12

FINAL REPORT

Predicting What People Learn from Examples

Richard Catrambone

May 1995

School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332-0170
USA

This research was sponsored by the Cognitive Science Program, Office of Naval Research, under Grant No. N00014-91-J-1137.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
5/11/95

3. REPORT TYPE AND DATES COVERED
Final Report 11/15/90 - 11/14/94

4. TITLE AND SUBTITLE
Predicting What People Learn from Examples

5. FUNDING NUMBERS
N00014-91-J-1137 G

6. AUTHOR(S)
Richard Catrambone

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
School of Psychology
Georgia Institute of Technology
Atlanta, GA 30332-0170

8. PERFORMING ORGANIZATION REPORT NUMBER
Deliverable #15

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Cognitive Science Program (1142CS)
800 N. Quincy Street
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Students often memorize a set of steps from examples in domains such as probability and physics, without inducing what subgoals those steps achieve. Such students fail to solve novel problems with identical goal structures but which do not permit exactly the same set of steps as the examples. This final report contains three papers that examine whether examples can help people to learn relevant subgoals for solving problems in a particular domain, and if learning the subgoals helps them to solve novel problems that involve those subgoals but require new steps for achieving them. It was hypothesized that when learners are encouraged to group steps from example solutions then they will be more likely to learn subgoals, perhaps through a self-explanation process. The connection between grouping and subgoal formation was supported by transfer results as well as analyses of participants' descriptions of how to solve problems. More generally, the fact that subgoals can be effectively conveyed by examples, and that these subgoals can aid transfer, has important implications for the design of training materials and tutoring environments.

14. SUBJECT TERMS
Problem solving, transfer, learning from examples

15. NUMBER OF PAGES
75

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE - See authorities.
NASA - See Handbook NHB 2200.2.
NTIS - Leave blank.

Block 12b. Distribution Code.

DOD - Leave blank.
DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA - Leave blank.
NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

★U.S.GPO:1991-0-296-362

FINAL TECHNICAL REPORT

Grant No: N00014-91-J-1137
Period: 11/15/90 - 11/14/94
Date of Submission: May 11, 1995
Name of Institution: Georgia Institute of Technology
Title of Project: Predicting What People Learn from Examples
Principal Investigator: Richard Catrambone

Table of Contents

Project Summary

(Summary of research carried out under ONR Grant No. N00014-91-J-1137.)

This document serves as the final technical report of ONR Grant No. N00014-91-J-1137. During this project I examined what people learn from examples in domains that emphasize solving quantitative problems. These domains would be areas such as algebra, probability, and physics. A central focus of this work was to develop a scheme for usefully representing the problem solving knowledge that one might acquire for a given domain. This was achieved using the notion of subgoals, which, in the present work, are used to represent the task structure for problems in a particular domain.

Subgoals show the higher-level structure that can be applied to a variety of problems that differ at superficial levels as well as in the mathematical details needed to solve the problems. Learners frequently focus on superficial aspects of examples and problems. This focus often prevents these learners from solving novel problems, that is, problems that cannot be solved by the same sequence of steps as in the training examples/problems. However, if a useful subgoal structure can be identified by a researcher or teacher, and if this structure can be conveyed to a learner, then the learner will be more successful solving novel problems. This implies that a crucial early step in teaching problem-solving in a domain is for teachers to spend time identifying to themselves the useful subgoals from a *novice's* perspective. One way this can be done is to first identify a target set of problems that the instructor wants the students to be able to solve. The instructor can then write out solutions to these problems and analyze them to determine the subgoals achieved by groups of steps that constitute the solution procedures to the problems.

A second focus of the research has been to find ways of effectively conveying subgoals to learners through examples once those subgoals have been identified by the instructor. The most effective technique found in the research has been to use manipulations that isolate groups of steps in the examples that achieve a particular subgoal. This isolation appears to serve as a cue to a learner that the steps go together. It is hypothesized that the learner then attempts to self-explain *why* the steps go together. The result of such a self-explanation process is a formation of a subgoal which can then be used in subsequent problems (even if the exact steps for achieving the subgoal are different in those problems compared to the example).

The project results suggest that training materials in problem-oriented domains should be designed to emphasize subgoal learning. This approach could improve learners' transfer to novel problems and situations. The project has also led to new work (with Department of Defense collaborators) on the application of the subgoal approach to automated (computer) training and tutoring environments.

The project has yielded a number of journal publications, proceedings papers, and conference presentations that are listed following this summary.

Project Publications and Reports

Journal Publications

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87* (1), 5-17.

Catrambone, R. (in press). Following instructions: The effects of principles and examples. *Journal of Experimental Psychology: Applied.*

Catrambone, R., Jones, C., Jonides, J., & Seifert, C. (1995). Reasoning about curvilinear motion: Using principles or analogy. *Memory & Cognition, 23* (3), 368-373.

Catrambone, R. (1994). Improving examples to improve transfer to novel problems. *Memory & Cognition, 22* (5), 606-615.

Papers in Refereed Conference Proceedings

Catrambone, R. (1995). Effects of background on subgoal learning. To appear in *Proceedings of the 17th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Catrambone, R. (1994). Cognitive science meets cognitive engineering. Symposium in *Proceedings of the 16th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum, 968-972.

Catrambone, R. (1994). The effects of labels in examples on problem solving transfer. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum, 159-164.

Catrambone, R. & Wachman. R.M. (1992). The interaction of principles and examples in instructions. In *Proceedings of the 14th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum, 749-754.

Catrambone, R. (1991). Helping learners acquire subgoals to improve transfer. In *Proceedings of the 13th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum, 352-357.

Paper Presentations

Catrambone, R. (1995). Effects of background on subgoal learning. Paper to be presented at the 17th Annual Conference of the Cognitive Science Society, Pittsburgh.

Catrambone, R. (1994). *Cognitive science meets cognitive engineering.* Symposium presented at the 16th Annual Conference of the Cognitive Science Society, Atlanta.

Catrambone, R. (1994). *The effects of labels in examples on problem solving transfer.* Paper presented at the 16th Annual Conference of the Cognitive Science Society, Atlanta.

Catrambone, R. (1994). *The effects of labels on subgoal learning.* Paper presented at the 35th Annual Meeting of the Psychonomic Society, St. Louis.

Catrambone, R. (1994). *Learning subgoals to solve problems.* Paper presented at the Office of Naval Research Grantees Meeting on Learning and Training, Chicago.

Catrambone, R. (1994). *What in the heck are subgoals?* Paper presented at the Fifth Annual Winter Text Conference, Jackson Hole, Wyoming.

Catrambone, R. (1993). *Conveying subgoals/subtasks to aid performance on novel problems.* Paper presented at the Fourth Annual Winter Text Conference, Jackson Hole, Wyoming.

Catrambone, R. (1993). *Effects of learning equations on transfer in mathematical domains.* Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, D.C.

Catrambone, R. (1992). *Improving examples to improve transfer.* Paper presented at the Third Annual Winter Text Conference, Jackson Hole, Wyoming.

Catrambone, R. (1992). *Producing better problem solving by making better examples.* Paper presented at the Lilly Endowment Teaching Fellows Conference, Callaway Gardens, Georgia.

Catrambone, R., Jonides, J., & Jones, C.M. (1992). *Reasoning about curvilinear motion: Using principles or analogy.* Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis.

Catrambone, R. (1991). *The effects of labels on learning subgoals for solving problems.* Paper presented at the 1991 Annual Meeting of the American Educational Research Association, Chicago, Illinois.

Catrambone, R. (1991). *Helping learners acquire subgoals to improve transfer.* Paper presented at the 13th Annual Conference of the Cognitive Science Society, Chicago.

Catrambone, R. (1991). *Learning from problem-solving examples.* Paper presented at the Office of Naval Research Grantees Meeting on Learning and Instruction, Atlanta.

Catrambone, R. (1991). *Subgoal learning not aided by studying multiple methods.* Paper presented at the 32nd Annual Meeting of the Psychonomic Society, San Francisco.

Catrambone, R. (1991). *Using labels to help learners acquire subgoals for solving novel problems.* Paper presented at the Fourth European Conference for Research on Learning and Instruction, Turku, Finland.

Invited Colloquia

Problem Solving and Transfer.
University of New England, March 1995.

Learning Subgoals from Examples to Solve Problems.
College of Staten Island, New York, December 1994.
Oklahoma State University, February 1995.
Wright State University, February 1995.

Improving Transfer to Novel Problems: The Use of Labels in Examples.
Armstrong Laboratory, Human Resources Directorate, Brooks Air Force Base, March 1994.

Memory & Cognition
1994, 22 (5), 606–615

Improving examples to improve transfer to novel problems

RICHARD CATRAMBONE
Georgia Institute of Technology, Atlanta, Georgia

People often memorize a set of steps for solving problems when they study worked-out examples in domains such as math and physics without learning what domain-relevant subgoals or subtasks these steps achieve. As a result, they have trouble solving novel problems that contain the same structural elements but require different, lower-level steps. In three experiments, subjects who studied example solutions that emphasized a needed subgoal were more likely to solve novel problems that required a new approach for achieving this subgoal than were subjects who did not learn this subgoal. This result suggests that research aimed at determining the factors that influence subgoal learning may be valuable in improving transfer from examples to novel problems.

A number of studies have indicated that learners rely heavily on worked-out examples when trying to solve novel problems (e.g., LeFevre & Dixon, 1986; Pirolli & Anderson, 1985). Unfortunately, in a domain, novices have great difficulty separating the features of the examples that are necessary to the solution procedure from those that are incidental (Ross, 1987, 1989). In addition, learners have difficulty generalizing solutions from examples to structurally similar, but nonisomorphic, problems (Reed, Ackinclose, & Voss, 1990; Reed, Dempster, & Ettinger, 1985). Although certain training manipulations have succeeded in improving transfer from examples to novel problems to some degree (Lewis & Anderson, 1985; Zhu & Simon, 1987), in general, transfer has not been impressive (e.g., Gick & Holyoak, 1983; Reed et al., 1985; Ross, 1987, 1989).

Learners differ in what they extract from worked-out examples. Chi, Bassok, Lewis, Reimann, and Glaser (1989) found that good learners are more likely to try to understand *why* a particular step was taken in a solution. Good and poor learners, at least initially after studying a physics chapter without examples, seemed to have a similar level of declarative knowledge about mechanics. However, when studying an example, the good learners produced explanations that contained more "inferences about the conditions, the consequences, the goals, and the meaning of various mathematical actions described in the example" (p. 168). Thus, good learners seem to get more from examples, including a knowledge of goals, even when starting at a knowledge level similar to that of poor learners.

The present experiments explore whether more learners can be turned into good learners by presenting them with examples that convey the subgoals relevant for solving problems in a domain. The term *subgoal* is used here to represent the task structure to be learned for solving problems in a particular domain (e.g., Catrambone & Holyoak, 1990; Dixon, 1987; Eylon & Reif, 1984). A subgoal groups a set of steps under a meaningful task or purpose (e.g., Anzai & Simon, 1979; Chi & VanLehn, 1991). I hypothesize that a person who has learned the subgoal will be in a better position to achieve it in a novel problem requiring a new or modified set of steps than will someone who has not learned the subgoal.

Learning subgoals is assumed to enhance performance because subgoals act as guides to the part or parts of the procedure demonstrated in examples that need to be changed for the current problem. Thus, subgoals narrow the space in which the solver has to search in order to determine what must be changed. For instance, Simon and Reed (1976) found that providing learners with a subgoal—in the form of a hint to achieve a particular state along the solution path in a problem involving missionaries and cannibals—aided their navigation through the problem space.

In the probability materials used in the first two experiments, one subgoal toward the overall goal of finding a particular probability is to find the probability of each of the individual events. If the steps for finding an individual event probability in a novel problem are not the same as those used in the example, then a person who has learned the subgoal to find the individual event probabilities will have a better chance of focusing on the steps of the procedure that must be changed—the steps involved in finding the individual event probabilities—than a person who has learned only a set of steps for finding the overall probability. For this second learner,

This research was supported by Office of Naval Research Grant N00014-91-J-1137. I thank Alana Anoskey and Elinor Nixon for their help in collecting and coding some of the data. I thank Dedre Gentner, Marsha Lovett, Brian Ross, Tim Salthouse, and Neff Walker for their comments on earlier drafts of this paper. Experiment 1 was reported at the Fourth Annual Winter Text Conference, January 1993, Jackson, Wyoming. Address correspondence to R. Catrambone, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332.

the steps for finding the individual probabilities are obscured because they are simply part of a longer set of steps for reaching the end goal. This learner would have fewer cues to direct him or her to the appropriate steps that need to be changed.

There may not be a theoretically "best" set of subgoals for solving problems in a domain. The particular subgoals that are taught might represent an instructor's judgment about how students should decompose problems into subproblems in order to solve novel problems most effectively. The judiciousness of the instructor's choice of subgoals can be measured by the success of the learners on novel problems.

A learner will be more likely to learn subgoals for the subparts of a general solution procedure if those subparts are emphasized in the examples' solutions. This claim hinges on the assumption that people will form a representation that is based on the most salient features of the example solution. The salience of a feature will vary, depending on the learner's expertise in the domain and how the solution is presented (Larkin, McDermott, Simon, & Simon, 1980; Ward & Sweller, 1990). In the present experiments, steps were labeled and visually separated as a technique for encouraging the formation of a subgoal.

An important first step in creating useful examples is to perform a task analysis to determine what elements need to be learned in order to solve problems in the domain of interest. *How* a domain is analyzed to produce those elements is by no means standardized. One promising approach is to create a set of production rules that solves problems or carries out procedures that one wants learners to be able to solve or learn (e.g., Anderson, Boyle, Farrell, & Reiser, 1987; Kieras & Bovair, 1986; Zhu & Simon, 1987). Examples can then be created in order to teach these productions to learners.

It does not seem necessary to make a commitment to a production rule formalism embodying a particular learning theory such as ACT* (Anderson, 1983) or Soar (Laird, Newell, & Rosenbloom, 1987) in order to derive the elements that need to be learned. However, a fundamental feature of most production rule systems—the goal structure—does provide a useful way to represent the knowledge needed to solve problems in a domain. Subgoals show the breakdown of a problem-solving procedure into subproblems (Anzai & Simon, 1979). Depending on the features of the examples and of the learner, the subgoals learned from examples could represent either a flexible and general approach to solving problems in a particular domain, or a rigid and superficial approach.

Learning From Examples

A number of studies have shown that manipulations of examples have a powerful and systematic effect on performance on novel problems (e.g., Catrambone & Holyoak, 1990; LeFevre & Dixon, 1986; Pirolli & Anderson, 1985; Reder, Charney, & Morgan, 1986; Ross, 1984). Given the central role that examples play in problem solving, and given the assumption that people learn subgoals from the examples, it is important to investigate the conditions that influence subgoal learning.

One reason that many learners do not form the "right" subgoals (as determined by an experimenter or instructor) is because examples typically are not designed to convey them. This observation echoes an inadequacy in a mechanics example from a physics textbook noted by Chi et al. (1989, p. 149). In the example, a block is suspended from a ceiling by two pieces of rope joined at a knot and a third piece of rope extending from the knot to the block. The task is to find the magnitude of two of the forces, given the third force. The solution states that the knot where the three strings are joined should be considered the body. However, no explanation is given as to why this decision is made. The decision is made because, in order to find a force in terms of other forces, the forces must all act on a common point. In this problem, the only place where all three forces act is the knot. This critical subgoal of finding a common point where the forces are acting would be useful for many future problems. However, instead of conveying this subgoal, the example is more likely to convey a series of steps that may or may not be useful for other problems.

One question at this point: Why not directly state the subgoals to learners rather than embedding them in examples? There are two problems with this approach. First, learners exhibit a clear preference for learning from and referring to examples when faced with new problems (e.g., LeFevre & Dixon, 1986; Pirolli & Anderson, 1985). Second, although there have been a small number of successes teaching solution procedures directly (Fong, Krantz, & Nisbett, 1986), most attempts have been unsuccessful (e.g., Reed & Bolstad, 1991).

Overview of Experiments 1 and 2

In the first two experiments, subjects studied examples that differed in whether they emphasized a subgoal that was predicted to be useful for solving novel test problems. It was predicted that a subject would be more successful at an unfamiliar part of a test problem if he or she had learned the relevant subgoal, compared with a subject who did not learn that subgoal.

The domain explored in Experiments 1 and 2 was probability. This domain was chosen because the training and testing materials can be relatively simple, and because prior work (Ross, 1987, 1989) has provided a useful manipulation.

Ross (1989; Experiment 1B) had subjects study example probability problems, such as ones involving permutations, and then solve several test problems. The mathematical roles of the entities (e.g., scientists, computers) in the examples and test problems were manipulated. For instance, Table 1 presents a permutation problem involving the determination of the probability that scientists will pick particular computers. The equation used for this example was $p = 1/[n(n-1)...(n-r+1)]$, where n is the number of choices available, and r is the number of choices being made. The test problems required some subjects to find, for instance, the probability that students will pick particular cars (humans picking inanimate objects), while other subjects had to find the probability that particular students would be assigned to particular cars (i.e., objects "picking" humans).

Corresponding mathematical roles are held by the humans and objects in the example in Table 1 and the test problem involving students picking cars. In both the example and the test problem, it is the number of humans, scientists and students, doing the choosing that provides the value for r. The number of objects, computers and cars, from which to choose provides the value for n. The second type of test problem, cars "picking" students, however, has reversed object correspondences: The humans provide the value for n (see Table 2B for another example of humans providing a value for n). Ross (1989) found that subjects were more successful at solving the first type of test problem than the second, presumably because their problem solving was guided to some degree by a feature correspondence approach. Specifically, if the number of objects provided the value of n in the examples, then subjects were likely to assign them this role in a test problem, even if it was the number of humans that should have provided this value in the test problem.

Besides working on permutation problems, Ross's subjects studied and solved combination problems. An example combination problem might ask for the probability that the seven hooks nearest the classroom door

Table 1
Permutation Training Example

The supply department at IBM has to make sure that scientists get computers. Today, they have 11 IBM computers and 8 IBM scientists requesting computers. The scientists randomly choose their computer, but do so in alphabetical order. What is the probability that the first 3 scientists, alphabetically, will get the lowest, second lowest, and third lowest serial numbers, respectively, on their computers?

Table 2
Test Problems

A. Permutation: People Picking Objects

As part of a new management policy, the Campbell Company is allowing the 20 company-owned vacation cottages to be used for vacations by their 14 plant managers. If the managers, in order of seniority, randomly choose a cottage from a list, what is the probability that the 4 managers with the most seniority get the most lavish, second most lavish, third most lavish, and fourth most lavish cottages, respectively?

B. Permutation: Objects Picking People

The secretaries at city hall are supposed to get new chairs this week. Today, city hall received 14 new chairs, and there are 11 secretaries requesting them. For inventory purposes, the property manager wants to assign the chairs in the order that they are unpacked. So, starting with the chair that is unpacked first, she randomly chooses a secretary to receive it, and continues until all the secretaries have chairs. What is the probability that the first 2 secretaries, alphabetically, will get the first and second chairs that are unpacked, respectively?

C. Combination: People Picking Objects

The Happy House Nursery School has had 17 hooks put up in the hall for the coats of their 14 students, with each student using 1 hook. The students each choose a hook at random as they come in one morning. What is the probability that the 7 tallest students get the 7 hooks closest to the classroom door? (It does not matter which of the particular 7 hooks closest to the door these students get, just as long as it is any 1 of the 7 closest.)

D. Combination: Objects Picking People

The Nashville Gnats Baseball team has a bus that has 30 seats. There are 25 players that are going on a road trip to play in a nearby town. To avoid arguments, the manager randomly chooses a player for each seat, starting with the seats in the front. What is the probability that the 6 pitchers get the 6 front seats? (It does not matter which of the particular 6 front seats the pitchers get, just as long as it is any 1 of the 6 in the front.)

Table 3
Solution Types Used for the Permutation Example in Table 1

Subgoal Solution

The equation needed for this problem is $1/[n*(n-1)*\ldots*(n-r+1)]$. In this problem, $n = 11$ and $r = 3$. However, another way of approaching the problem is to think of it in the following way:

Probability of the first scientist (who comes first alphabetically) getting the computer with the lowest serial number = 1/11.

Probability of the second scientist getting the second lowest serial number = 1/10.

Probability of the third scientist getting the third lowest serial number = 1/9.

So, $1/11*1/10*1/9 = 1/990$ = overall probability.

Equation Solution

The equation needed for this problem is $1/n*(n-1)*\ldots*(n-r+1)$. This equation allows one to determine the probability of the above outcome occurring. In this problem, $n = 11$ and $r = 3$. The 11 represents the number of computers that are available to be chosen, and the 3 represents the number of choices that are being focused on in this problem. The equation divides the number of ways the desired outcome could occur by the number of possible outcomes. So, inserting 11 and 3 into the equation, we find that $1/11*10*9 = 1/990$ = overall probability.

would be picked by the seven tallest students in a class (see Table 2C). The equation used to solve combination problems of this sort is $p = [h!(j-h)!]/j!$, where h is the number of entities (e.g., students) doing the choosing, and j is the number of entities in the pool from which things are chosen (e.g., hooks). Again, Ross demonstrated the object correspondence phenomenon.

Although Ross taught his subjects the procedures for solving both permutation and combination problems and examined transfer to problems in which the roles of humans and objects were switched, an examination of the two procedures shows that a more general procedure can be used to solve both problem types. Both permutation and combination problems can be analyzed by considering the individual event probabilities that contribute to an overall probability. This approach is demonstrated in the "subgoal" solution provided for the problem of the scientists and the computers (see Table 3). The combination problem, involving students and coathooks (Table 2C), can be analyzed in a similar way:

Probability that one of the seven tallest students will get a hook near the door = 7/17.
Probability that one of the remaining six tallest students will get a hook near the door = 6/16.
Probability that one of the remaining five tallest students will get a hook near the door = 5/15, etc.
So,

$$\frac{7}{17} * \frac{6}{16} * \frac{5}{15} * \frac{4}{14} * \frac{3}{13} * \frac{2}{12} * \frac{1}{11}$$

$$= \frac{7!}{17 * 16 * \ldots * 11} = \text{overall probability.}$$

Combination problems have numerators that are no longer simply "1." Instead, they start at the number of acceptable choices and then are decremented just like the denominator.

The "subgoal" solution, presented in Table 3 for the scientists and computers permutation problem, is assumed to help learners form two goals. The first goal is to find the overall probability; this goal is assumed to be formed because it is explicitly stated in the example. The second is the subgoal to find each event probability—for example, the probability that the first scientist will get the computer with the lowest serial number, the probability that the second scientist will get the computer with the second lowest serial number, and so on. This subgoal is assumed to be formed because each individual event probability is explicitly labeled and spatially separate in the subgoal solution in Table 3. The method for finding an individual event probability will involve the steps of inserting a 1 in the numerator and placing the number of (remaining) objects in the denominator of each probability.

The "equation" solution, presented in Table 3 for the scientists and computers permutation problem, is assumed to help learners form only the goal to find the overall probability. The method for achieving this goal will consist of a set of steps for finding numbers from the problem statement and inserting them into the equation.

Performance Predictions

Subjects who study either the equation solution or the subgoal solution are predicted to perform well on the first test problem they are given (see Table 2A). This first problem is a permutation problem that is isomorphic to the training examples that the subjects studied, and humans and objects play the same roles that were in the examples (i.e., humans picking objects). The subjects can simply repeat the steps learned from the examples. Performance on this problem serves as a check that a subject has learned at least a set of steps.

Two of the three remaining test problems, one permutation problem and one combination problem (see Tables 2B and 2D), reverse the roles for humans and objects compared with the training examples; that is, objects are "picking" humans. What this means in terms of the solution is that the numbers that go into the denominator are based on the number of humans in the problem, not the number of objects. It is predicted that subjects will have difficulty with this aspect of the problems. However, those who study the subgoal solution (the subgoal group) are predicted to have learned the subgoal to find each individual event probability, and thus might be more likely to consider what the numerator and denominator mean in each event probability. As a result, these subjects have a better chance to consider modifying the denominator for this problem than subjects in the equation group.

The two combination test problems (see Tables 2C and 2D) are expected to cause difficulty on numerator performance since, unlike the examples' numerators, the numerators are no longer 1. Once again, the Subgoal group, by virtue of learning the subgoal to find each individual event probability, is predicted to outperform the equation group because the subgoal of finding each individual event probability might lead the Subgoal subjects to consider the numerator as a potential locus for change.

EXPERIMENT 1

In Experiment 1, I tested the hypothesis that transfer will be improved if subjects study training examples that emphasize a subgoal needed for novel problems. Besides examining problem-solving performance, another measure of subgoal learning was attempted by having subjects describe how to solve probability problems after studying the examples. If a subject's description includes a statement such as "find the probability of each event's occurrence," then this would be taken as additional support that he or she had learned that subgoal. The subgoal group should mention the subgoal of breaking the overall probability into a set of individual event probabilities more often than the equation group, since this is the major difference in the solution approaches presented to the two groups. The subjects who mention this subgoal should perform better than other subjects on the denominators for reversed-roles permutation and combination problems and on the numerators for the combination problems.

Although the emphasis so far has been placed on the differences in the solutions studied by the subgoal and equation groups, it is important to note that the solutions are similar in length and, presumably, clarity (see Table 3). One could construct a solution type that is arbitrarily unclear and demonstrate that subjects' performance on transfer problems is poor relative to subjects who study examples using a solution type that is arbitrarily clear. It is suggested that the equation solution contains potentially useful information for solving test problems, but that subjects who study that solution are likely to focus on how values are inserted into the equation rather than form a more general procedure.

Method

Subjects. The subjects were 66 students from introductory psychology classes at the Georgia Institute of Technology who participated for course credit. None of the subjects had taken a probability course prior to participating in the experiment.

Materials and Procedure. The subjects received a booklet containing training examples and test problems. All the subjects studied two isomorphic worked-out permutation example problems in which humans picked objects. Table 1 presents one of these examples.

The subjects were randomly assigned to the equation group (n = 31) or the subgoal group (n = 35). In the equation group's examples, an equation was used in order to solve the problem. The solution included an explanation of the meaning of the numbers being inserted into the equation. The subgoal group's examples divided the problem into finding each individual probability. Table 3 contains the solutions that were seen by the groups for one of the training examples.

The subjects were asked to study the examples carefully and were told that, after studying them, they would be asked to solve some problems. They were also told that they could not look at the examples when working on the problems. This restriction was intended to increase the likelihood that they would pay attention to the examples and how they were solved.

After studying the examples, the subjects were asked to describe how to solve problems in the domain. The instructions were: "Suppose you were going to teach someone how to solve probability problems of the type you have just studied. Please describe the procedure you would give someone to solve these problems. Please be as complete as possible."

After writing their descriptions, the subjects attempted to solve the four test problems in Table 2. The first problem was isomorphic to the examples (see Table 2A). The second problem (B) was a permutation problem like the examples, but humans provided the value of n (see Table 2B). The third and fourth problems (C and D) were combination problems. In the first combination problem, humans picked objects, which is the same notion shown in the examples (see Table 2C). In the second combination problem, objects picked humans (see Table 2D). Thus, the test problems represented a range of difficulty as a function of whether they involved permutations or combinations and whether humans were picking objects, or objects were picking humans.

The subjects worked at their own pace and were asked to show all their work. In general, they took about 30 min to complete the experiment. Each permutation problem was scored for whether a subject used the correct denominator. For instance, the solution to the second permutation problem is 1/11 * 1/10. If a subject wrote 1/14 * 1/13, confusing the roles of the chairs and secretaries, the denominator would be scored as incorrect. For combination problems, the numerator and denominator were both scored as correct or incorrect. Two raters independently scored the problems; their scores agreed 92% of the time. Any disagreements were resolved by discussion. The frequencies with which the groups correctly found the denominators for the permutation problems and the denominators and numerators for the combination problems were analyzed by using the likelihood ratio chi-square test (G^2; Bishop, Fienberg, & Holland, 1975).

Results and Discussion

The overall performance differences between the groups can be summarized as follows: They did not reliably differ on denominator performance for the reversed-role problems (i.e., objects picking humans), but the subgoal group outperformed the equation group on the numerators for the combination problems.

As expected, both groups were quite successful in determining the denominator on the first permutation problem, which was isomorphic to the training examples and had humans and objects playing the same roles as in the examples [$G^2(1) = 0.5, p = .78$; see Table 4].

In the second permutation problem, objects picked humans—a reversal from the training examples. An error that the subjects frequently made on this problem was to use the number of chairs (14) as the starting point in the denominator rather than the number of secretaries (11). Although the groups did not significantly differ in finding the denominator [$G^2(1) = 1.11, p = .29$], there was an 11% advantage for the subgoal group (see Table 4).

The next problem was a combination problem in which humans picked objects. The subjects were fairly successful at finding the denominator and, as expected, did not differ significantly on this measure [$G^2(1) = .09$, $p = .76$; see Table 4]. However, they did have difficulty finding the correct numerator for this problem; many of the subjects simply used 1. As expected, the subgoal group outperformed the equation group on this measure [$G^2(1) = 6.39, p = .01$].

In the second combination problem, objects picked humans. As in the reversed-roles permutation problem, the subjects had difficulty finding the correct value for the denominator. The groups did not differ significantly on this measure [$G^2(1) = .06, p = .80$; see Table 4]. As in the first combination problem, the subjects had difficulty finding the correct numerator. As expected, the subgoal group was more successful than the equation group on this measure [$G^2(1) = 9.83, p = .002$].

Table 4
Performance (Percent Correct) on Experiment 1 Test Problems

	Group	
	Equation (n = 31)	Subgoal (n = 35)
Permutation Problem 1 (people pick objects)		
Denominator	94	89
Permutation Problem 2 (objects pick people)		
Denominator	23	34
Combination Problem 1 (people pick objects)		
Denominator	71	74
Numerator	13	40
Combination Problem 2 (objects pick people)		
Denominator	23	20
Numerator	10	43

The transfer results suggest that the subjects in both groups were equally misled by superficial role reversals of objects and humans in the denominator. However, the subgoal subjects were more likely to adapt their solution procedure to find the numerator correctly in combination problems.

Relationship between training examples and explanations produced by subjects on how to solve problems. The subjects' explanations of how to solve problems in the domain were scored for whether they mentioned the subgoal of dividing the overall probability into a series of individual probabilities. Six equation subjects and 5 subgoal subjects produced explanations that were too general or idiosyncratic to be scored. These explanations typically consisted of statements such as "I would read through the example and write it up on the board for the person." These subjects were excluded from the following analyses.

The subgoal group mentioned the notion of dividing the overall probability into a set of individual probabilities far more often than the equation group [83% vs. 8%; $G^2(1) = 35.3, p < .0001$]. This was expected, since the subgoal group studied example solutions that labeled and isolated individual probabilities, whereas the equation group did not.

Relationship between explanations and transfer performance. It was expected that the subjects who mentioned the subgoal of finding individual event probabilities in their explanations would be more likely to correctly find the denominator in the reversed-role problems and the numerator in the combination problems.

There was no difference in denominator performance for the first permutation problem for the subjects who mentioned the notion of dividing a probability into individual probabilities (the "IndProb" subjects) compared with the subjects who did not mention this notion in their explanations (the "OneProb" subjects) [$G^2(1) = .18, p = .67$; see Table 5]. This is not surprising, since in this problem humans choose objects, as in the examples.

Table 5
Performance (Percent Correct) on Test Problems as a Function of Subjects' Explanations in Experiment 1

	Group	
	OneProb* ($n = 27$)	IndProb† ($n = 28$)
Permutation Problem 1		
Denominator	89	93
Permutation Problem 2		
Denominator	14	37
Combination Problem 1		
Denominator	78	74
Numerator	14	44
Combination Problem 2		
Denominator	11	26
Numerator	11	44

*Does not mention breaking problem into individual probabilities.
†Mentions breaking problem into individual probabilities.

As expected, the IndProb subjects were more successful than the OneProb subjects at finding the correct value for the denominator in the second permutation problem, which involved reversed roles for humans and objects [$G^2(1) = 3.84, p = .05$; see Table 5].

In the first combination problem, humans and objects played the same roles that were in the examples, and therefore it is not surprising to find that there was no difference between the groups in finding the correct value for the denominator in this problem [$G^2(1) = .15, p = .69$; see Table 5]. As expected, the IndProb subjects were more successful than the OneProb subjects at finding the correct value for the numerator [$G^2(1) = 6.26, p = .01$; see Table 5].

Although the anticipated difference between the groups in finding the correct value for the denominator in the second combination problem—a reversed-roles problem—was not found [$G^2(1) = 2.18, p = .14$], the IndProb group had a 15% advantage (see Table 5). As expected, the IndProb subjects were more successful than the OneProb subjects at finding the correct numerator value [$G^2(1) = 8.29, p = .004$; see Table 5].

These results follow the trend of those that were found when the instructional groups were compared on the transfer problems. This makes sense, since the subgoal subjects were by far the ones most likely to mention the notion of dividing an overall probability into individual probabilities.

It could be argued that learners who write "better" descriptions (e.g., mention the subgoal of finding individual event probabilities) are also the ones who are better at transfer. One defense against this argument is to note that the experimental manipulation of type of examples studied influenced transfer success.

Although the subgoal subjects did not reliably outperform the equation subjects on the denominators for reversed-role problems, they were clearly superior at finding the correct numerators for the combination problems. In addition, the IndProb group outperformed the OneProb group on finding the denominator for one of the reversed-role problems. Although most of the results support the predictions, the unreliability of the denominator effect for the reversed-role problems suggested that a second, more focused experiment would be appropriate.

EXPERIMENT 2

The procedure and materials for Experiment 2 were identical to those of Experiment 1, except for three features: (1) the subjects studied three rather than two training examples, (2) the subjects were not asked to write explanations of how to solve problems, and (3) for half of the subjects, the combination problems did *not* contain the last sentence shown for each combination problem in Table 2.

The number of examples presented to the subjects was increased to improve the likelihood that they would learn the procedures demonstrated in the examples and perhaps improve transfer to the test problems. They were not asked to write explanations, since this task was unusual and may have influenced learning in unanticipated ways. Finally, the last sentence for each combination problem in Table 2 was removed for half of the subjects in order to examine whether they signaled to the subjects that these problems were different from the training examples and needed to be approached differently. Perhaps with this cue removed, the subgoal subjects would not outperform the equation subjects.

Method

Subjects. The subjects were 78 students from introductory psychology classes at the Georgia Institute of Technology who participated in the experiment for course credit. None of the subjects had taken a probability course prior to participating in the experiment.

Materials and Procedure. The subjects received a booklet containing training examples and test problems. They all studied three isomorphic worked-out permutation example problems in which humans picked objects (including the two used in Experiment 1).

The subjects were randomly assigned to the equation group ($n = 40$) or the subgoal group ($n = 38$). After studying the examples, they solved the four test problems in Table 2. The subjects worked at their own pace and were asked to show all their work. In general, they took about 25 min to complete the experiment. Two raters independently scored the problems, and their scores agreed 90% of the time. Any disagreements were resolved by discussion.

Table 6
Performance (Percent Correct) on Experiment 2 Test Problems

	Group	
	Equation ($n = 40$)	Subgoal ($n = 38$)
Permutation Problem 1		
Denominator	90	95
Permutation Problem 2		
Denominator	22	42
Combination Problem 1		
Denominator	85	84
Numerator	0	29
Combination Problem 2		
Denominator	30	32
Numerator	0	32

Results and Discussion

The results were similar to those from Experiment 1 and supported most of the predictions. The subgoal group strongly outperformed the equation group on the numerators for the combination problems and showed a trend toward superior performance on the denominator for one of the two reversed-role problems.

As expected, both of the groups did well in determining the denominator for the first permutation problem, and there was no significant difference in their performance [$G^2(1) = 0.63, p = .43$; see Table 6]. As in Experiment 1, both of the groups showed inferior performance in determining the denominator for the second permutation problem—a problem with reversed roles for humans and objects. As predicted, there was an advantage (20%) for the subgoal group, although this difference just missed significance [$G^2(1) = 3.47, p = .06$; see Table 6].

Performance on the combination problems was initially broken down as a function of training group and whether or not the problems contained the last sentence presented for each combination problem in Table 2. There was no main effect of sentence for either problem, nor was there an interaction between group and presence of the sentence; thus, the analyses and the results for the combination problems are collapsed across this factor in Table 6.

In the first combination problem, humans picked objects, and, as anticipated, the subjects were fairly, and equally, successful at finding the denominator [$G^2(1) = .009, p = .92$; see Table 6]. As expected, the subgoal subjects outperformed the equation subjects at finding the correct numerator [$G^2(1) = 17.7, p = .0001$; see Table 6].

The predicted superior performance by the subgoal subjects did not occur for finding the correct value for the denominator in the second combination problem, a reversed-roles problem [$G^2(1) = 0.02, p = .88$; see Table 6]. As expected, the subgoal subjects outperformed the equation subjects at finding the correct numerator [$G^2(1) = 19.6, p = .0001$; see Table 6].

It was hypothesized that the subgoal subjects learned the subgoal to find individual event probabilities, but the equation subjects did not. This subgoal was predicted to aid performance in finding the denominator for reversed-role problems and the numerators for combination problems. Both experiments clearly demonstrated the numerator effect. Experiment 1 demonstrated a trend toward the denominator effect for the reversed-role problems, primarily when the subjects were partitioned into those who mentioned, or failed to mention, the subgoal of finding individual event probabilities in their explanations. This trend was stronger for the permutation problem than for the combination problem. Experiment 2 also showed this effect more strongly for the permutation problem. It is not clear if there are certain features of these two problems that differentially affected denominator performance. The subjects' explanations of how to solve problems were consistent with the claim that the subgoal subjects were more likely to learn the subgoal to find individual event probabilities than the equation subjects.

The fact that the denominator effect was less reliable than the numerator effect may be due to the role of superficial features. The training examples presumably led the subjects to expect objects to provide the value for the denominator, and perhaps this expectation tended to override any benefits due to learning subgoals, especially since a value for objects was always provided in each problem. The tendency to put a 1 as the numerator may have been more easily overridden by subgoal learning because its connection with either humans or objects is less clear.

EXPERIMENT 3

Experiment 3 was an attempt to generalize the findings from the first two experiments to another domain: algebra word problems. This domain was chosen because prior work had demonstrated poor transfer from training to transfer problems despite attempts to improve examples (e.g., Reed et al., 1985). Consider the algebra example in Table 7A, in which one has to determine how long it would take someone to do a job given that certain information about their work rate and time and another person's work rate and time are provided. This problem involves using an equation for determining work that requires representing each worker's work rate and time: $(\text{rate}_1 \times \text{time}_1) + (\text{rate}_2 \times \text{time}_2) = 1$.

Learners are good at memorizing how to solve problems that are isomorphic to the one in Table 7A. In this problem, both of the workers' rates are represented as constants. The time spent working by Worker 1 is represented as a variable, and Worker 2's time is represented as a function of that variable. However, learners may not encode the example solution by determining a representation for each rate and time and then inserting these representations into the equation. Instead, they have a more superficial understanding of the solution procedure, which involves matching the form used in the example, finding similar values in the problem statement, and inserting them into the equation. As a result, if a new problem requires a different representation of the rates and times, these learners may be unable to solve the problem. That is, the learners may not have learned that certain subgoals exist—the subgoals of representing each worker's rate and time—and that these subgoals might be achieved differently (i.e., different ways of representing rate and time depending on the givens in the problem) from the way they were achieved in the example.

For instance, the problem in Table 7B requires that Worker 1's rate be represented as a variable. In addition, instead of having the workers' times be represented as a variable and a function of that variable, the times are represented as a constant and a function of that constant. Nevertheless, the new representations can be inserted into the same equation that was used for the example in Table 7A. Similarly, the problem in Table 7C requires that Worker 1's rate be represented as a variable, and Worker 2's rate be a function of that variable. Their times are both represented as constants. These representations are different from those used in the examples.

It is hypothesized that if the representations for rates and times are highlighted separately from the equation in the example solutions, then learners will be more likely to learn that rate and time are individual representations that must be determined for each worker. In addition, it is hypothesized that if subjects learn the subgoals of representing workers' rates and times, then they will be more likely to correctly solve a novel problem requiring novel representations for rates and times than would subjects who do not learn those subgoals.

Table 7
Example and Test Work Problems in Experiment 3

A. Mary can rebuild a carburetor in 3 hours and Mike can rebuild one in 4 hours. How long would it take Mary to rebuild a carburetor if she and Mike work together, but Mike works for ½ hour more than Mary?

Solution
- $1/3$ = Mary's rate
- t = time Mary spent rebuilding carburetor
- $1/4$ = Mike's rate
- $t + 1/2$ = time Mike spent rebuilding carburetor
- $(1/3 * t) + [1/4 * (t + 1/2)] = 1$
- $7/12 * t = 1 - 1/8$
- $t = 7/8 * 12/7 = 3/2$ hours = time Mary spent rebuilding carburetor

B. Mr. Jones can refinish a dresser in 5 hours. After working for 2 hours, he is joined by his wife. Together they finish the job in 1 hour. How much of the job could his wife do in 1 hour when working alone?

Solution [not seen by subjects]
- $[1/5 * (2+1)] + (w * 1) = 1$
- $3/5 + w = 1$
- $w = 2/5$ = wife's rate
- so, in 1 hour, wife could do 2/5 of job

C. Barbara and Connie can finish a job in 6 hours when they work together. Barbara works twice as fast as Connie. How much of the job could Connie do in 1 hour when working alone?

Solution [not seen by subjects]
- $(2c * 6) + (c * 6) = 1$
- $12c + 6c = 1$
- $18c = 1$
- $c = 1/18$ = Connie's rate
- so, in 1 hour, Connie could do 1/18 of job

Subjects who learn the subgoal to represent each worker's rate should be more successful at representing the rate as a variable in the problem in Table 7B and rate as a variable and rate as a function of a variable in the problem in Table 7C. Subjects who learn the subgoal to represent each worker's time should be more successful at representing the time as a constant and a function of a constant in the second problem (B) and as a constant in the third problem (C).

Method

Subjects. Sixty-two students from introductory psychology classes at the Georgia Institute of Technology participated for course credit.

Materials and Procedure. The subjects studied three isomorphic example word problems dealing with work, including the example in Table 7A. The "rate and time label" (RTL) group (n = 21) studied examples in which the representations for rates and times were presented separately from the equation (see Lines 1–4 in the solution to the example in Table 7A). The "time label" (TL) group (n = 20) studied examples that presented the representations for each worker's time (i.e., Lines 2 and 4 from the example in Table 7A). The "rate label" (RL) group (n = 21) studied examples that presented the representations for each worker's rate (i.e., Lines 1 and 3 from the example in Table 7A).

After studying the examples, the subjects received three problems to solve. One was isomorphic to the training examples, and the other two involved new and old ways of representing rate and/or time (see Tables 7B and 7C).

Results and Discussion

All of the groups performed well at representing rate and time in Problem 1, which was isomorphic to the

Table 8
Performance (Percent Correct) on Test Problems in Experiment 3

	Group		
	RTL ($n = 21$)	TL ($n = 20$)	RL ($n = 21$)
Problem 1			
Rate Worker 1 (constant)	100	100	100
Rate Worker 2 (constant)	100	100	95
Time Worker 1 (variable)	100	100	100
Time Worker 2 (function of a variable)	95	95	100
Problem 2			
Rate Worker 1 (constant)	100	95	100
Rate Worker 2 (variable)	86	55	86
Time Worker 1 (function of a constant)	81	90	57
Time Worker 2 (constant)	90	90	57
Problem 3			
Rate Worker 1 (variable)	90	60	90
Rate Worker 2 (function of a variable)	86	50	86
Time Worker 1 (constant)	90	100	62
Time Worker 2 (constant)	90	100	62

Note—RTL, rate and time label; TL, time label; RL, rate label.

training examples (see Table 8). This is not surprising, since the subjects could match the representations from the training examples and simply plug in the new values.

In Problem 2 (Table 8), almost all of the subjects correctly represented the rate as a constant for Worker 1. Again, this is reasonable since this was the representation used in the training examples. The RTL and RL groups were significantly more successful than the TL group at representing the rate for Worker 2 as a variable [$G^2(2) = 6.63, p = .04$]. The RTL and TL groups were more successful than the RL group at representing the rate for Worker 2 as a constant [$G^2(2) = 8.78, p = .01$] and Worker 1 as a function of that constant [$G^2(2) = 6.50, p = .04$].

In Problem 3 (Table 8), the RTL and RL groups were significantly more successful than the TL group at representing the rate for Worker 1 as a variable [$G^2(2) = 7.59, p = .02$] and Worker 2 as a function of that variable [$G^2(2) = 8.63, p = .01$]. The RTL and TL groups were more successful than the RL group at representing the rates for Workers 1 and 2 as constants [$G^2(2) = 13.66, p = .001$, in both cases]. Across the problems, the most common errors were that the subjects either left out the representation for time or rate in the equation, or wrote that not enough information was given in the problem.

The results from this experiment are consistent with the hypothesis that subjects who learned the subgoals of representing workers' rates and times would represent them more successfully on novel problems. However, alternate explanations exist. One is that the highlighting manipulation essentially provided labels for the variables, and thus made them more meaningful. This "meaningfulness" helped the subjects to properly use the variables in the equation for the novel transfer problems. The results from Experiment 3, taken alone, are not sufficient to discriminate between this explanation and the subgoal explanation. However, the pattern of results across the three experiments is consistent with a subgoal explanation. It is not clear how the "subgoal" solution in the first two experiments made the variables more meaningful.

One could consider the subgoal explanation of the results as being an attempt to make the notion of meaningfulness more precise. Subgoals can be viewed as a way of making a solution procedure more meaningful by providing guideposts that the solver must reach en route to achieving the overall solution to the problem. In the algebra experiment, the guideposts were to explicitly represent the rate and time for each worker. Even if these guideposts are somewhat arbitrary, at least from the point of view of the solver, they still provide organization and guidance that may make the solver less likely to stray from the correct solution path (cf. Mawer & Sweller, 1982).

It is worth noting that performance on the test problems was generally quite good. This suggests that even for relatively sophisticated subjects (most of the students in the experiment had at least one term of college calculus), examples that more effectively convey subgoals can improve transfer performance.

GENERAL DISCUSSION

The aim of the present study was to examine whether examples that teach a subgoal structure for solving problems in a domain could be created and whether learning these subgoals would help subjects solve problems that required novel methods for them. The numerator performance results and some of the denominator performance results in the probability experiments, and the rate and time results in the algebra experiment, suggest that subgoals can be conveyed to learners through examples, and that learning these subgoals helps people achieve them in novel problems. This is quite encouraging in comparision with the usual finding in the problem-solving literature, which has shown poor transfer from training materials to test problems that require more than a simple repetition of a set of memorized steps (e.g., Reed et al., 1985; Ross, 1987, 1989). These findings suggest that examples that emphasize a useful subgoal structure can help turn learners into the "good" learners observed by Chi et al. (1989), who tended to find meaning, such as goals, for the mathematical steps in the examples and who made use of this information when solving novel problems.

Mayer and Greeno (1972) experimentally manipulated the meaningfulness of instruction in solving binomial probability problems by varying whether the instruction focused on mechanical operations or on concepts that were presumed to be part of subjects' prior knowledge. They found that the "mechanical" group was more successful at solving familiar problems, whereas the "concept" group was better at answering "understanding" questions about the domain, such as whether the number of successes could be greater than the number trials. The present study extends Mayer and

Greeno's findings by showing that learning subgoals promotes transfer to novel problems while also helping learners exhibit some level of understanding, as shown by the explanations produced by the subgoal subjects in Experiment 1.

CONCLUSIONS

The present results are consistent with the claim that learning subgoals will help learners determine which parts of a solution procedure need to be modified in order to solve novel problems. The results also provide a starting point for determining how to "emphasize" a subgoal in an example. The approach used here was to create example solutions that isolated components of the procedure that could be construed as subgoals. The visual separation and the labeling, and perhaps their interaction, may have all played a role in subgoal learning.

Smith and Goodman (1984) examined subgoal learning by comparing a group of subjects who followed a set of steps for assembling an electric circuit with a group who received a structurally oriented "explanatory schema" with the steps. This schema consisted of statements that provided a rationale for carrying out sets of steps. Each rationale was essentially a statement of a goal that the steps were achieving (e.g., "The next thing that you will have to do is to assemble the on–off switch"). When assembling a new circuit, the subjects who had previously received the explanatory schema were more accurate at building the substructures corresponding to the goals, even though the required steps were not identical to the ones followed during training.

The present results, and those from Smith and Goodman (1984), suggest that research aimed at determining factors that affect subgoal learning, such as the use of labels and visual separation of steps in examples, would have clear pedagogical benefits. In addition, it is important, both in terms of theory development and the production of effective training materials, to explore what constitutes a good subgoal structure for a given domain. Perhaps a theory-motivated technique can eventually be developed for determining effective subgoal structures for any given domain.

REFERENCES

ANDERSON, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

ANDERSON, J. R., BOYLE, C. F., FARRELL, R., & REISER, B. J. (1987). Cognitive principles in the design of computer tutors. In P. Morris (Ed.), *Modeling cognition*. New York: Wiley.

ANZAI, Y., & SIMON, H. A. (1979). The theory of learning by doing. *Psychological Review*, **86**, 124-140.

BISHOP, Y. M. M., FIENBERG, S. E., & HOLLAND, P. W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: MIT Press.

CATRAMBONE, R., & HOLYOAK, K. J. (1990). Learning subgoals and methods for solving probability problems. *Memory & Cognition*, **18**, 593-603.

CHI, M. T. H., BASSOK, M., LEWIS, R., REIMANN, P., & GLASER, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, **13**, 145-182.

CHI, M. T. H., & VANLEHN, K. (1991). The content of physics self-explanations. *Journal of the Learning Sciences*, **1**, 69-106.

DIXON, P. (1987). The processing of organizational and component step information in written directions. *Journal of Memory & Language*, **26**, 24-35.

EYLON, B., & REIF, F. (1984). Effects of knowledge organization on task performance. *Cognition & Instruction*, **1**, 5-44.

FONG, G. T., KRANTZ, D. H., & NISBETT, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, **18**, 253-292.

GICK, M. L., & HOLYOAK, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, **15**, 1-38.

KIERAS, D. E., & BOVAIR, S. (1986). The acquisition of procedures from text: A production system analysis of transfer of training. *Journal of Memory & Language*, **25**, 507-524.

LAIRD, J. E., NEWELL, A., & ROSENBLOOM, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, **33**, 1-64.

LARKIN, J., MCDERMOTT, J., SIMON, D. P., & SIMON, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, **208**, 1335-1342.

LEFEVRE, J., & DIXON, P. (1986). Do written instructions need examples? *Cognition & Instruction*, **3**, 1-30.

LEWIS, M. W., & ANDERSON, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology*, **17**, 26-65.

MAWER, R. F., & SWELLER, J. (1982). Effects of subgoal density and location on learning during problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **8**, 252-259.

MAYER, R. E., & GREENO, J. G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology*, **63**, 165-173.

PIROLLI, P. L., & ANDERSON, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skills. *Canadian Journal of Psychology*, **39**, 240-272.

REDER, L. M., CHARNEY, D. H., & MORGAN, K. I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory & Cognition*, **14**, 64-78.

REED, S. K., ACKINCLOSE, C. C., & VOSS, A. A. (1990). Selecting analogous problems: Similarity versus inclusiveness. *Memory & Cognition*, **18**, 83-98.

REED, S. K., & BOLSTAD, C. A. (1991). Use of examples and procedures in problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 753-766.

REED, S. K., DEMPSTER, A., & ETTINGER, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 106-125.

ROSS, B. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, **16**, 371-416.

ROSS, B. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 629-639.

ROSS, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 456-468.

SIMON, H. A., & REED, S. K. (1976). Modeling strategy shifts in a problem-solving task. *Cognitive Psychology*, **8**, 86-97.

SMITH, E. E., & GOODMAN, L. (1984). Understanding written instructions: The role of an explanatory schema. *Cognition & Instruction*, **1**, 359-396.

WARD, M., & SWELLER, J. (1990). Structuring effective worked examples. *Cognition & Instruction*, **7**, 1-39.

ZHU, X., & SIMON, H. A. (1987). Learning mathematics from examples and by doing. *Cognition & Instruction*, **4**, 137-166.

(Manuscript received July 7, 1993;
revision accepted for publication October 29, 1993.)

Journal of Educational Psychology
1995, Vol. 87, No. 1, 5–17

0022-0663/95/$3.00

Aiding Subgoal Learning: Effects on Transfer

Richard Catrambone
Georgia Institute of Technology

Students often memorize a set of steps from examples in domains such as probability and physics, without learning what subgoals those steps achieve. A result of this sort of learning can be that these students fail to solve novel problems that do not permit exactly the same set of steps even though the old goal structure is maintained. Three experiments demonstrated that both labeling and visually isolating a set of steps in examples independently help students learn a subgoal and be more likely to solve novel problems that involve that subgoal but require different steps to achieve it.

What do students learn from worked examples in domains such as probability and physics? Often, students memorize a set of steps to which they attach relatively little meaning. As a result, when faced with a test problem that has the same goal structure but requires different steps for achieving the goals, students may rigidly and incorrectly apply the old steps or fail to produce a solution at all (e.g., Reed, Dempster, & Ettinger, 1985; Ross, 1987, 1989).

Learners are more successful solving novel problems when they learn the goal structure of the problems in that domain (e.g., Anzai & Simon, 1979; Eylon & Reif, 1984). Researchers use the term *subgoal* (and *goal*) in two different ways. The first defines a subgoal to be something people—or computer programs—form when they are working on a problem and reach a point where they do not simply recognize what to do next because they have no options, too many options, conflicting options, etc. A subgoal is formed at this impasse (e.g., Newell, 1990, Chapter 4; VanLehn, 1988). The second considers subgoals to represent the task structure to be learned for solving problems in a particular domain and assumes that these subgoals can be taught to learners (e.g., Catrambone, 1994a; Catrambone & Holyoak, 1990; Dixon, 1987; Eylon & Reif, 1984). From the second point of view, a subgoal groups a set of steps under a meaningful task or purpose (e.g., Anzai & Simon, 1979; Chi & VanLehn, 1991). For instance, in the probability materials used in the current experiments, a set of multiplication and addition steps can be grouped under the subgoal "find the total frequency of the event." It is the second, task-analysis-driven view of subgoals that I followed in the present study.

This research was supported by Office of Naval Research Grant N00014-91-J-1137. Experiment 1 was reported at the Third Annual Winter Text Conference in Jackson Hole, Wyoming, in January 1992.

I thank Susan Bovair, Janet Kolodner, Margaret Recker, Tim Salthouse, Kimberly Sessions, John Sweller, and Barry Gholson for helpful comments on drafts of this article. I thank Alana Anoskey, Margaret Dasher, and Paul Gay for their help in collecting the data and Elinor Nixon for her help in coding the data.

Correspondence concerning this article should be addressed to Richard Catrambone, School of Psychology, Georgia Institute of Technology, Atlanta, Georgia 30332. Electronic mail may be sent via Internet to rc7@prism.gatech.edu.

An instructor might help students learn the goal structure through worked examples. The particular subgoals taught might represent an instructor's judgment about how students should decompose problems into subproblems to solve novel problems effectively. Novel problems are taken to be problems that share the same goal or task structure with training examples but that require a change in the steps for achieving at least one of the subgoals.

Much research indicates that the way learners encode examples influences how likely they are to access examples and how likely they are to apply or adapt examples successfully in new situations (e.g., Brown, Kane, & Echols, 1986; Gentner & Gentner, 1983; Gick & Holyoak, 1983). For instance, Brown et al. found that young children were more likely to use a prior story to help them solve an analogous problem if the children had either spontaneously, or through prompting, previously induced the goal structure of the story.

My aim in this study was to examine whether learners' transfer to novel problems is improved if they study examples designed to convey a solution procedure organized by subgoals versus examples designed to convey only the steps of the solution procedure.

Learning From Examples

Learners, particularly novices, typically prefer to study or refer to examples, as opposed to instructions or descriptions of principles of a domain, when working on problems (e.g., Pirolli & Anderson, 1985). Learners explicitly mention examples when solving a problem (Lancaster & Kolodner, 1988; Ross, 1984), and they follow examples rather than instructions when the two conflict (LeFevre & Dixon, 1986).

Unfortunately, people frequently do not learn from examples what is needed to solve novel problems. Rather, people tend to memorize a set of steps. Attempts to improve this situation have usually found little improvement in transfer. For instance, Reed et al. (1985) provided elaboration designed to help students understand the principles illustrated in algebra examples. In general, these elaborations did not improve performance on nonisomorphic problems. Similar transfer difficulties after learning from examples were documented by Gick and Holyoak (1983), Ross (1987, 1989), and others. However, several researchers have demonstrated some success in using examples to help learners transfer to novel problems (e.g., Ward & Sweller, 1990; Zhu & Simon, 1987). It is unclear why conflicting transfer results are found across these studies. The subgoal framework offered below may help to identify situations in which transfer is more or less likely to occur.

Value of Learning Subgoals

Prior work suggests that students can learn subgoals from examples. Chi and her colleagues (Chi & Bassok, 1989; Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi & VanLehn, 1991) examined the self-explanations learners produce when studying examples. Chi et al. (1989) divided their students into "good and poor students" (p. 158) as a function of test problem performance that followed an example-studying phase. "Good" students demonstrated superior transfer to novel problems; their self-explanations of the steps from example solutions in physics mechanics problems contained more goals, descriptions of the preconditions for actions, and explications of the consequences of actions, than did the self-explanations of "poor" students (Chi & VanLehn, 1991).

Although students who learn easily and students who have difficulty learning may differ on various dimensions in addition to the self-explanations they produce, one inference that might be drawn from Chi and her colleagues' work is that learning will be improved if examples provide the types of explanations that students who learn easily produce on their own. One of these explanation types are subgoals.

Subgoals are useful because they group a set of steps and in some sense explain what the steps accomplish. If the learner can recognize which subgoals are relevant for solving a novel problem, then those subgoals can guide the learner to the steps from the old solution procedure that need to be modified to achieve the subgoals in the new problem.[1] In contrast, a learner who had simply memorized a string of steps for solving a particular problem type, without grouping sets of steps under the subgoals they achieve, will have fewer cues to direct him or her to the steps that need to be modified for the novel problem.

A subgoal framework can be useful for explaining transfer. The framework can be used to determine the subgoals learners need to know to solve problems in a domain, and it can be used to guide the construction of examples to increase the likelihood that learners learn these subgoals. Researchers conducting studies within this framework can examine issues such as finding the most effective ways to determine the subgoals in a domain and finding effective ways of conveying those subgoals through examples.

Improving Examples

In the current study, a set of steps, which was part of the overall solution, was labeled in training examples as a way of conveying the subgoal those steps achieved. I hypothesized that example solutions that label a set of steps will make a person more likely to learn the subgoal achieved by the steps rather than merely memorizing the steps themselves. Furthermore, I hypothesized that by learning the subgoal, the learner will be more likely to successfully achieve it in a novel problem that requires a different set of steps to achieve that subgoal.

This second hypothesis was made for two reasons. First, subgoals were one type of knowledge produced more often by students who learned easily than by students who had difficulty learning in Chi and her colleagues' studies (e.g., Chi et al., 1989). Second, subgoal learning seems to be correlated with better performance at solving problems that involve the same subgoals, even if the problems are novel and require different steps to achieve the subgoals (Anzai & Simon, 1979; Catrambone & Holyoak, 1990; Mawer & Sweller, 1982).

The present study continues a line of research on factors that influence subgoal learning (Catrambone, 1994a; Catrambone & Holyoak, 1990). Catrambone and Holyoak (1990) examined whether learners studying examples that demonstrated multiple ways of achieving a subgoal would be more likely to learn the subgoal than learners studying examples demonstrating a single method. The results suggested that studying multiple methods did not aid subgoal learning, at least within the confines of a 1-hr experiment. However, subgoal learning did appear to be aided when learners studied examples that provided elaborations about the particular methods used for achieving a subgoal. However, these elaborations contained a variety of information including additional domain theory, explanations of the conditions that led to a certain method being chosen, and labels for key groups of steps. Thus, it is not clear which feature(s) of the elaboration promoted subgoal learning.

On the basis of preliminary research (Catrambone, 1994b) and the findings of Chi et al. (1989) suggesting that students who learn easily tend to group steps into subgoals, I examined whether a relatively minimal manipulation, providing a label for a set of steps, would make learners more likely to form a subgoal.

A study by Smith and Goodman (1984) supports the idea that labels can aid subgoal learning and transfer to new situations. The relevant comparison is between a group of students who followed a set of steps for assembling an electric circuit and a group who received a structurally oriented "explanatory schema" (Smith & Goodman, 1984, p. 360) with the steps. This schema consisted of statements that provided a rationale for why sets of steps needed to be

[1] The issue of learning to recognize when a subgoal is appropriate for a particular problem is not addressed here. All training and test problems involved the same subgoals. Rather, the emphasis is on learners' ability to achieve the subgoal in a new way in test problems.

carried out. Each rationale was essentially a statement of a goal that the steps achieved (e.g., "The next thing that you will have to do is to assemble the on–off switch."). When assembling a new circuit, students who had previously received the explanatory schema were more accurate in building the substructures corresponding to the goals even though the required steps were not identical to the ones followed during training.

The current study differs from Smith and Goodman's (1984) work in several ways. First, during training their students carried out a set of instructions, whereas students in the current experiments studied examples. Second, the Smith and Goodman transfer task involved the students' following a new set of instructions, whereas in the present study students had to solve novel problems. Third, the explanatory schema used by Smith and Goodman provided a rationale for why a set of steps were carried out. Conversely, although the labels used in the present experiments presumably helped students group a set of steps, it was left to the students to supply the rationale for why those steps were executed. Thus, in the present study I analyzed transfer in a more demanding situation in which I explicitly examined the role of examples in subgoal learning.

Test Domain: The Poisson Distribution

The Poisson distribution is often used to approximate binomial probabilities for events occurring with some small probability. The Poisson equation is

$$P(X = x) = \frac{[(e^{-\lambda})(\lambda^{x})]}{x!},$$

where λ is the average (the expected value) of the random variable X, e is the mathematical constant 2.718, and x is the number of successes of interest.

Consider an example of the use of the Poisson distribution in which the average number of briefcases owned per lawyer is found so that one can predict the probability of a randomly chosen lawyer owning a certain number of briefcases (see Appendix A). The method for the goal of finding λ, the average frequency of the event (e.g., owning a briefcase), could be represented as:

Goal: Find λ

Method:

1. Multiply each event category (e.g., owning exactly zero briefcases, owning exactly one briefcase, etc.) by its observed frequency.
2. Sum the results.
3. Divide the sum by the total number of trials (number of lawyers) to find the average number of briefcases per lawyer.

Learners are good at memorizing, from examples, the above steps to achieve the subgoal of finding the average frequency of an event for problems with different storylines (Catrambone & Holyoak, 1990). However, they often fail to notice that Steps 1 and 2 in the above method could also be viewed as a method for achieving the subgoal of finding the total frequency of the event (e.g., total number of briefcases owned). As a result, learners often have trouble finding the average frequency of an event when a problem provides the total frequency of that event directly (see Appendix B), rather than requiring that it be derived from the frequencies of various event-categories (Catrambone, 1994b; Catrambone & Holyoak, 1990).

If a person learns the subgoal "find the total frequency of the event," he or she might be better able to find λ in a novel problem that requires a change from the examples in how total frequency is found. For example, in the method discussed earlier for finding the average number of briefcases owned per lawyer, it might be better if the learner's method for finding λ was organized as follows:

Goal: Find λ

Method:

1. Goal: Find total number of briefcases.

   Method:

   a. Multiply each event-category by its observed frequency.
   b. Sum the results to obtain the total number of briefcases.

2. Divide the total number of briefcases by the total number of trials to obtain the average number of briefcases per lawyer.

Experiment 1

In Experiment 1, I examined subgoal learning by manipulating whether a set of steps was labeled. It was assumed that if learners see a label for a set of steps, they are more likely to link those steps to a common subgoal. Subgoal learning was then assessed in two ways. In the first, I analyzed transfer performance—how successfully students found λ—on novel problems by students who were hypothesized to have learned or not to have learned the subgoal of finding the total frequency of an event. In the second, I had these students describe how to solve problems in the Poisson domain. If the descriptions of students in the label condition included statements such as "find the total frequency" and if these were the students who also solved the novel problems more successfully, then this would provide converging evidence that these students learned that subgoal.

In this experiment the no-label group studied examples demonstrating the weighted average method for finding λ (see Appendix A, the no-label solution). The label group's examples differed in that the steps for finding the total frequency were explicitly labeled rather than merged with the overall set of steps for finding λ (see Appendix A, the label solution).

In the test phase students were asked to (a) describe how they would teach someone to solve problems like the ones they had studied and (b) solve problems requiring the use of the weighted average method to find λ or a method in which the total frequency was supplied directly (see Appendix B).

Predictions

It was hypothesized that the label group would be more likely than the no-label group to to learn the subgoal to find the total frequency. As a result, students in the label group should be more likely than students in the no-label group to mention the idea of finding a total frequency in their descriptions. It could be argued that this is a trivial prediction because the label group could simply be repeating a label or a generalization of a label from the examples. Nevertheless, if the students who mention the idea of finding total frequency are primarily those in the label group and if these students produce superior transfer performance on problems that provide the total frequency directly, then support would be given to the claims that labels aid subgoal learning and that subgoals aid transfer.

No difference was predicted between the groups in the frequency of mentioning the notion of finding an average (or λ) in their descriptions because all examples mentioned the term *average* in the solutions.

Both groups were predicted to be quite successful solving transfer problems that were isomorphic to the training examples. The label group was expected to perform better than the no-label group on test problems that provided the total frequency directly rather than requiring that it be calculated.

Method

Participants. Participants were 48 students recruited from an introductory psychology class at the Georgia Institute of Technology who received course credit for their participation. None of the students had taken a probability course before participating in the experiment.

Materials and procedure. All students initially studied a cover sheet that briefly described the Poisson distribution and how it could be used as a replacement for more cumbersome techniques for calculating probabilities involving events that can be categorized as successes and failures. The Poisson equation was presented, and a simplified notion of a random variable was also presented.

Students were randomly assigned to one of two groups. The label group (N = 25) studied six examples demonstrating the weighted average method for finding λ in which the steps for finding the total frequency were explicitly labeled (see the label solution in Appendix A for an example). The no-label group's (N = 23) examples differed in that the steps for finding the total frequency were not labeled (see the no-label solution in Appendix A). The labels seen by the students in the label group were specific to the context of the problem (e.g., "total number of briefcases owned") and were not phrased at a general level (e.g., "total frequency") so that I could minimize additional explicit general domain instruction.

After studying the examples, students were asked to describe how to solve problems in the domain. The instructions were:

> Suppose you were going to teach someone how to solve Poisson distribution problems of the types you have just studied. Please describe the procedure or procedures you would give someone to solve these problems. Please be as complete as possible. Please do not look back at the examples.

After writing their descriptions, students solved three test problems that required the use of the weighted average method for finding λ (isomorphic to the example in Appendix A), and then they solved three test problems in which total frequency was given directly (and thus λ could be found by simply dividing the given number by the total number of trials). The latter type included the problem in Appendix B and two problems isomorphic to it. Students were told not to look back at the examples when solving the test problems.

Students' written solutions were scored for whether they found λ correctly. In addition, students' descriptions of how to solve problems in the domain were scored for two primary features: an explicit mention of trying to find the total frequency and an explicit mention of trying to find an average. Two raters independently scored the descriptions and problem solutions and agreed on scoring 94% of the time. Disagreements were resolved by discussion.

Results

Students were given a score of 1 for a given problem if they found λ correctly and a score of 0 otherwise. The scores for the three problems that were isomorphic to the training examples, Problems 1–3, were summed, thus giving students a score ranging from 0 to 3 for their performance on those problems. Similarly, the scores for the three novel problems, Problems 4–6, were summed, thus giving students a score ranging from 0 to 3 for their performance on those problems.

Transfer as a function of group. As expected, both groups did quite well at finding λ on test problems that were isomorphs to the training examples (Problems 1–3), and there was no significant difference in performance, $F(1, 46) = 0.28$, $p = .60$, $MSE = .40$. The average for the label group was 2.9, and the average for the no-label group was 2.8.

As predicted, however, the label group found λ more successfully than did the no-label group on the three test problems that involved the new method for finding λ (Problems 4–6), $F(1, 46) = 5.32$, $p = .03$, $MSE = 1.82$. The average for the label group was 2.2, and the average for the no-label group was 1.3. The most frequent mistake that students made on these problems was to write in the solution area that not enough information was given to solve the problem.

Descriptions as a function of group. Students' descriptions were scored for whether they explicitly mentioned finding the total frequency and the subgoal of finding the average or λ. As expected, the groups were equally likely to mention finding the average (label, 60% and no-label, 65%), $\chi^2(1) = 0.14$, $p = .71$. As predicted, the label group mentioned finding total frequency more often than did the no-label group (52% vs. 13%), $\chi^2(1) = 8.18$, $p = .004$.

Transfer as a function of descriptions. Students who mentioned finding the total frequency in their descriptions (N = 16) tended to perform about the same on the isomorphic test problems when compared with students who did not mention it (N = 32), $F(1, 46) = 1.70$, $p = .20$, $MSE =$

0.39 (Ms = 3.0 and 2.75, respectively). As predicted, students who mentioned finding the total frequency did better than did the other students at finding λ on test problems that required the new method for finding λ, $F(1, 46) = 5.52$, $p = .02$, $MSE = 1.81$ (Ms = 2.38 and 1.41, respectively).

Students who mentioned finding the average in their descriptions ($N = 30$) did not perform significantly differently on the isomorphic test problems when compared with students who did not mention it ($N = 18$), $F(1, 46) = 2.06$, $p = .16$, $MSE = 0.39$, (Ms = 2.93 and 2.67, respectively). There was also no significant difference between these groups at finding λ for the problems that required the new method for finding λ, $F(1, 46) = 0.20$, $p = .66$, $MSE = 2.02$ (Ms = 1.80 and 1.61, respectively).

Discussion

Students in the no-label group were not able to find λ as successfully in the novel problems (Problems 4–6) as were the students in the label group. This result suggests that students in the label group were more likely to learn the subgoal of finding the total frequency as part of the solution structure for solving Poisson distribution problems. As a result, the students in the label group were able to find λ in problems that could not be solved with the steps from the training examples.

Converging evidence that the students in the label group had learned the subgoal to find the total frequency comes from students' descriptions of how to solve problems. Students in the label group were four times more likely than students in the no-label group to mention the goal of finding the total frequency. Both groups mentioned the goal of finding λ equally often which was consistent with the fact that λ was labeled in the training examples for both groups.

It could be argued that those students who wrote descriptions in which they mentioned the subgoal of finding the total frequency were also people who were simply better at transfer. However, the finding that the label and no-label groups had differential far transfer success (i.e., success at solving problems requiring new or modified steps compared to the examples) suggests that the training manipulation affected subgoal learning in addition to any effects that were due to individual differences.

The results from this experiment are consistent with the claim that manipulations to examples that promote attention to subgoals can help students learn those subgoals. Students in the label group learned the subgoal to find the total frequency, and this subgoal helped their performance on test problems that required a new way to find the total frequency. However, this new way was to recognize that the total frequency was given in the problem. Perhaps a more stringent test of whether learning a subgoal helps a learner to achieve it when a novel method is required would be to give students test problems that require a new method of calculating total frequency. Experiment 2 was designed to do this.

Experiment 2

In Experiment 2, I examined students' ability to modify an old method for calculating the total frequency to find λ. As in Experiment 1, the steps for finding the total frequency were either labeled or unlabeled in the training examples.

The new method for finding total frequency was to add the frequencies of a number of events rather than to multiply event- categories by their frequencies and then add the products (such as the method used in Appendix A). One of the two test problems requiring this modified method involved children finding seashells on the beach (see Appendix C). The numbers of shells found by a child on Day 1, Day 2, Day 3, Day 4, and Day 5 are given. The problem then asks for the probability of a randomly chosen child finding a certain number of shells. If the label group is more likely than the no-label group to learn the subgoal of finding the total frequency, then the students in the label group should be more successful than the students in the no-label group at modifying their approach for finding total frequency and thus, finding λ in this problem.

The solutions presented for the training examples were the same as in Experiment 1 except that the steps for finding total frequency and finding λ were circled, either separately or together (see Figure 1). Students saw both presentations for each example and were asked to pick the one they felt circled the steps that "go together." It was hypothesized that if a student learned the subgoal to find the total frequency, then he or she would be more likely to prefer that the steps for finding the total frequency be separated from the step for finding λ. (This is the solution in Figure 1A for the label students and Figure 1C for the no-label students.)

For each test problem, students were asked to circle the parts of their solutions that went together. Again, because students in the label group were predicted to be more likely than students in the no-label group to have learned the subgoal to find the total frequency, it was predicted that students in the label group would be more likely than students in the no-label group to circle the steps for finding the total frequency separately from the step for finding λ. Students' circling performance could provide converging evidence, along with their transfer performance, for subgoal learning.

Method

Participants. Participants were 52 students from an introductory psychology class at the Georgia Institute of Technology who participated for course credit. None of the students had taken a probability course before participating in the experiment. Students were randomly assigned to the label ($N = 26$) and no-label ($N = 26$) groups.

Materials and procedure. Both groups studied the same cover sheet used in Experiment 1 and then studied three examples illustrating the weighed average method of finding λ. For the label group, the subgoal of finding the total frequency of the event was labeled. This subgoal was not labeled for the no-label group. The examples were a subset of those used in Experiment 1.

Students saw two solutions to each example. The solutions were identical except for how the steps were circled. For instance, for

A. Label Students, Circled Separately

Total number of briefcases owned = [1(180) + 2(17) + 3(13) + 4(9)] = 289

$$E(X) = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X=x) = \frac{[(e^{-\lambda})(\lambda^{x})]}{x!}\text{, so}\quad P(X=2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

B. Label Students, Circled Together

Total number of briefcases owned = [1(180) + 2(17) + 3(13) + 4(9)] = 289

$$E(X) = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X=x) = \frac{[(e^{-\lambda})(\lambda^{x})]}{x!}\text{, so}\quad P(X=2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

C. No-Label Students, Circled Separately

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X=x) = \frac{[(e^{-\lambda})(\lambda^{x})]}{x!}\text{, so}\quad P(X=2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

D. No-Label Students, Circled Together

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X=x) = \frac{[(e^{-\lambda})(\lambda x)]}{x!}\text{, so}\quad P(X=2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

Figure 1. Example solutions for the problem in Appendix A in which the steps for finding the total frequency and λ are circled either separately or together in Experiment 2.

the training example in Appendix A, students in the label group saw the solutions in Figures 1A and 1B, whereas students in the no-label group saw the solutions in Figures 1C and 1D. Students were given the following instructions:

> For each example, please note that the solution is presented twice. The two presentations of each solution have the steps circled. The circles are used to indicate steps that might "go together." For instance, suppose you were following a recipe for cooking something. Perhaps the first three steps of the recipe involved putting various ingredients into a bowl and the fourth step involved stirring the ingredients with a spoon and the fifth step involved using a blender to finishing the mixing. You might draw a circle around the first three steps because they involve "adding ingredients," and you might draw a circle around the the fourth and fifth steps because they involve "blending the ingredients." The two solutions for each example are identical except that different steps are in the circles. For each example, indicate whether Presentation 1 or Presentation 2 is the one with which you most agree.

As a counterbalancing measure, the solutions that had the steps for finding the total frequency and λ in the same circle came first for half of the students in each group, whereas for the rest of the students the solution that had the steps for finding the total frequency and finding λ circled separately came first. The solution order had no effect on students' preferences for solutions or on transfer performance, and thus, was collapsed over for all analyses.

After studying the examples, students solved five target problems. The first was a weighted average problem isomorphic to the examples. The second also involved finding λ as a weighted average; however, the divisor (total number of trials) was not given directly in the problem. Rather, it had to be found by adding the number of members in each category (see Appendix D). The third provided the value of the total frequency directly (see Appendix B). The fourth and fifth involved a modification of the old method for finding the total frequency: instead of the total frequency being found by multiplying the event-categories by their frequencies and then summing, it was found by adding a set of frequencies (see Appendix C for an example).

After they solved each test problem students were asked to circle the steps of their solution that went together. Students were told not to look back at the examples when working on the test problems.

Predictions

Training examples solution preferences. Students in the label group were predicted to be more likely than students in the no-label group to prefer solutions in which the steps for finding the total frequency were circled separately from the steps for finding λ. However, it should be noted that the "circled separately" presentation for the students in the no-label group (see Figure 1C) looks a bit odd because a set of steps that are part of a fraction are circled, and the denominator is left out of the circle. Thus, students could be predisposed not to choose this presentation.

Transfer performance. It was predicted that both groups would do well on the first test problem, an isomorph to the training examples. For the second problem both groups were expected to perform similarly. Given that the subgoal of finding the number of trials was not emphasized in the training examples for either group, performance was expected to be poor. As in Experiment 1, the label group was predicted to do better at finding λ on the third test problem, which involved the recognition that the total frequency was given directly. The label group was predicted to do better on the fourth and fifth test problems, which involved finding the total frequency by adding a set of frequencies. The subgoal of finding the total frequency should have aided students in the label group in figuring out which steps they needed to modify to find the total frequency.

Segmenting of solutions to test problems. Students in the label group were predicted to be more likely than students in the no-label group to circle the steps for finding the total frequency separately from the step for finding λ in their solutions to the test problems that involve calculating the total frequency (Problems 1, 2, 4, and 5). No circling difference was predicted between the groups for the problem in which the total frequency was given directly (Problem 3) because it would have been rather odd to circle the given total frequency as a step.

Two raters independently scored the example solution preferences and problem solutions, and they agreed 88% of the time. Disagreements were resolved by discussion.

Results and Discussion

Training examples solution preferences. A given student invariably chose the same circling scheme across the three examples (i.e., either the one in which the steps for finding total frequency and the step for finding λ were circled separately or the one in which they were circled together). This was not surprising because the training examples were isomorphs. Thus, students were categorized into one of two groups: those who preferred separate circles for total frequency and λ and those who preferred that total frequency and λ be in the same circle. As predicted, students in the label group were more likely than students in the no-label group to choose solutions in which the steps for finding the total frequency were circled separately from the step for finding λ (50% vs. 19%), $\chi^2(1) = 5.44$, $p = .02$.

Transfer performance. All but one student, from the label group, correctly solved the first test problem, a weighted average problem isomorphic to the training examples.

The second test problem was also a weighted average problem, but the total number of trials, the value that would be placed in the denominator when calculating λ, was not given directly. Rather, the student had to calculate the total number of trials by adding the number of workers. As expected, the proportion of students who found λ correctly in each group did not differ significantly (label, 85% and no label, 73%), $\chi^2(1) = 1.04$, $p = .31$, although the overall performance was higher than expected.

The third test problem provided the value for the total frequency directly. As predicted, the label group correctly found λ significantly more often than did the no-label group (88% vs. 54%), $\chi^2(1) = 7.59$, $p = .006$. This replicates the finding from Experiment 1.

The fourth and fifth test problems required the students to calculate the total frequency by adding the simple frequencies (e.g., the number of shells found each day). Students were given a score of 1 for a given problem if they found λ correctly and a score of 0 otherwise. The scores for the two problems were summed, thus giving students a score ranging from 0 to 2 for their performance on those problems.

As predicted, the label group found λ more successfully than did the no-label group, $F(1, 50) = 4.15$, $p = .047$, $MSE = 0.46$ (Ms = 1.50 and 1.12, respectively). Students who solved these problems incorrectly tended, for example, to multiply day by number of shells (Problem 4) or write that not enough information was given. This result suggests that learning a subgoal does aid a person in achieving it in a novel problem that requires a modification of an old method.

Segmenting of solutions to test problems. For each test problem that required λ to be calculated (Problems 1, 2, 4, and 5), students were given a score of 1 if they circled the steps for finding the total frequency separately from the step for finding λ in their solution and a score of 0 otherwise. The scores for the problems were summed, thus giving students a score ranging from 0 to 4.

As expected, students in the label group were more likely to circle the steps for finding the total frequency separately from the step for finding λ in their solutions than were the students in the no-label group, $F(1, 50) = 8.39$, $p = .006$, $MSE = 1.32$ (Ms = 1.15 and 0.23, respectively).

Also as expected, there was no difference between the groups on Problem 3 in which the total frequently was provided directly, $\chi^2(1) = 2.08$, $p = .15$. Only two students, both in the label group, circled the total frequency. The rest of the students simply circled the entire set of calculations that they used for finding λ (if they found a value for λ).

Experiment 3

The results from the first two experiments are consistent with the interpretation that the design of examples that promote attention to subgoals can be successful in helping students to learn those subgoals and to transfer more successfully. In both experiments a labeling manipulation was used to promote attention to the subgoals. However, the examples containing a label for total frequency differed from the no-label examples in an additional way: The steps for finding total frequency were on a separate line from the rest of the steps for finding λ. Thus, it is possible that this visual isolation either by itself or in combination with the label, produced the superior transfer. Perhaps manipulations that lead learners to group a set of steps will make those learners more likely to form a subgoal to relate those steps. In Experiment 3, I examined this possibility with four groups of students.

The visual-isolation group studied example solutions such as the one in Appendix E (this is a solution to the problem in Appendix A) in which the total frequency steps were on their own line without a label. The separate-line label group studied example solutions isomorphic to those seen by the label group in Experiment 1 (see Appendix A). Note, though, that this solution style, besides labeling the steps for finding the total frequency, places those steps on a separate line from the rest of the steps for finding λ. Thus, students in this group saw solutions that involved both labeling and visual isolation. To examine whether there was any interaction between labeling and visual isolation, I had the same-line label group study examples that had the steps for finding the total frequency labeled but had those steps located on the same line as the rest of the steps for finding λ (see Appendix E). Finally, a no-label group analogous to the group from Experiment 1 (see Appendix A) was included as a baseline.

According to the label view, the label is primarily responsible for subgoal learning. Thus, students who study examples in which the steps for finding total frequency are labeled—the separate-line label and the same-line label groups—should outperform the visual-isolation and no-label groups on transfer problems because the steps for finding total frequency were not labeled in the examples studied by the latter two groups.

According to the grouping view, grouping—and thus subgoal learning—is promoted by labeling or visual isolation. Under this view, the label groups and the visual-isolation group should perform similarly, and all these groups should outperform the no-label group. Neither view explicitly predicted an interaction between labeling and visual isolation. However, if a label and visual isolation are required for grouping and subgoal formation, then the separate-line label group should perform better than all other groups on the transfer problems.

Method

Participants. Participants were 118 students from an introductory psychology class at the Georgia Institute of Technology who participated for course credit. None of the students had taken a probability course before participating in the experiment.

Materials and procedure. All students studied the same cover sheet used in the prior experiments and then studied two isomorphic examples illustrating the weighted average method for finding λ. The examples were a subset of those used in Experiment 1. Fewer study examples were used in this experiment because pilot testing with students receiving six, three, or two examples showed no effect of number of examples on performance.

Students were randomly assigned to one of four groups. Students in the same-line label group ($N = 30$) had the steps for finding the total frequency labeled, but the steps were on the same line as the rest of the steps for finding λ. The solution to the problem in Appendix A studied by the same-line label group is shown in Appendix E. The separate-line-label group ($N = 30$) also had the steps for finding total frequency labeled, but they were on a separate line from the rest of the steps for finding λ. The solution to the problem in Appendix A that was studied by the separate-line label group was identical to the one studied by the label group in Experiment 1 (see Appendix A). The visual-isolation group ($N = 29$) studied examples in which the steps for finding total frequency were unlabeled but on their own line (see Appendix E). The no-label group ($N = 29$) saw solution types identical to those studied by the no-label group in Experiment 1 (see Appendix A).

As in Experiment 1, students were asked to describe how to solve problems in the domain after they finished studying the examples. After writing their descriptions, students solved three test problems. The first problem required the use of the weighted average method for finding λ (isomorphic to the example in Appendix A). The second problem provided the total frequency directly (see Appendix B). The third problem involved the addition of simple frequencies to get the total frequency (see Appendix C).

Students were told not to look back at the examples when solving the test problems.

Students' written solutions were scored for whether they found λ correctly. Students' descriptions of how to solve problems in the domain were scored for whether the students mentioned trying to find the total frequency and trying to find an average. Two raters independently scored the descriptions and problem solutions and agreed on scoring 90% of the time. Disagreements were resolved by discussion.

Results and Discussion

Transfer as a function of group. As expected, all groups did quite well at finding λ on the test problem that was isomorphic to the training examples. Only two students, one in the same-line label group and one in the separate-line label group, solved this problem incorrectly.

For the two far transfer problems, students were given a score of 1 for a given problem if they found λ correctly and a score of 0 otherwise. The scores for the two problems were summed, thus giving students a score ranging from 0 to 2 for their performance on those problems.

There were significant differences among the four groups with respect to finding λ in the novel test problems, $F(3, 114) = 5.53$, $p = .0014$, $MSE = 0.52$, with means of 1.7, 1.6, 1.5, and 1.0 for same-line label, separate-line label, visual-isolation, and no-label groups, respectively. As predicted by the grouping view, Shaffer (1986) sequential Bonferroni pairwise comparisons (familywise $\alpha = .05$) indicated that both label groups and the visual-isolation group outperformed the no-label group, all $ps < .0167$. Also consistent with the grouping view, but not with the label view, was the finding that no reliable performance differences were found between the two label groups or between either label group and the visual-isolation group (all $ps > .05$).

Descriptions as a function of group. Students' descriptions were scored for whether they explicitly mentioned the subgoal of finding the total frequency and the subgoal of finding the average or λ. As expected, the groups were equally likely to mention finding λ (same-line label, 87%; separate-line label, 87%; visual isolation, 76%; and no label, 83%), $\chi^2(3) = 1.62$, $p = .65$.

There were significant differences among the four groups with respect to the frequency of mentioning finding total frequency, $\chi^2(3) = 17.43$, $p = .0006$, with percentages of 43%, 57%, 31%, and 7% for same-line label, separate-line label, visual-isolation, and no-label groups, respectively. Consistent with the grouping view and the transfer results, pairwise comparisons (familywise $\alpha = .05$) indicated that the two label groups each mentioned finding total frequency more often than did the no-label group (both $ps < .0167$, as did the visual-isolation group ($p < .0167$, one-tailed). Differences among the separate-line, same-line, and visual-isolation conditions were not statistically significant according to the Shaffer (1986) sequential Bonferroni procedure, all $ps > .025$.

Transfer as a function of descriptions. As expected, there were no significant performance differences on the novel test problems between students who mentioned finding λ ($N = 98$) and those who did not ($N = 20$), $F(1, 116) = 1.04$, $p = .31$, $MSE = 0.57$ ($Ms = 1.5$ and 1.3, respectively).

Students who mentioned finding the total frequency ($N = 41$) were more successful than other participants ($N = 77$) at finding λ on the novel test problems, $F(1, 116) = 5.77$, $p = .018$, $MSE = .55$ ($Ms = 1.7$ and 1.3, respectively).

The results of Experiment 3—similar transfer performance by the label groups and visual-isolation groups, and better performance than the no-label group—support the grouping view of subgoal formation and also suggest that there is no apparent interaction between labeling and visual isolation on subgoal formation. The tendency of the label and visual-isolation groups to mention finding total frequency in their descriptions more often than the no-label group is also consistent with the claim that these students were more likely to form a subgoal for finding total frequency.

Finally, students who mentioned finding total frequency in their descriptions showed better transfer on the novel transfer problems. Students in Experiment 1 who mentioned total frequency also transferred better on the novel problems. In addition, in a prior study (Catrambone, 1994b) I also found this relationship between descriptions and transfer performance. Together, these studies suggest that students' descriptions provide a potentially useful measure of subgoal learning. The transfer and description results provide converging evidence for the hypothesis that manipulations that encourage grouping enhance subgoal learning and thus, transfer.

General Discussion

Learners' problem-solving knowledge in domains such as probability often seems to be focused on the mathematical steps illustrated in training examples. This translates into poor performance on transfer problems. If subgoals could be conveyed to learners, then learners should be more successful on novel problems. One way these subgoals can aid learners is by guiding them to the steps in the old solution procedure that need to be changed to achieve the subgoals in a novel problem.

All three experiments presented here suggest that if people are led to learn a subgoal, in this case either by studying examples that label the steps that achieve the subgoal or by visually isolating those steps, then they are more likely to successfully achieve it in a novel problem. It is impressive to find a performance difference between the label and visual-isolation groups compared with the no-label groups given that the new methods for finding the total frequency in the test problems were seemingly straightforward: adding a set of frequencies or recognizing that the total frequency was given directly. This may partly be an effect that is due to *Einstellung* or mechanization (Luchins, 1942): Students may have been biased against altering a memorized set of steps.

Besides transfer differences, the experiments provide additional evidence that manipulations that encourage grouping can help people learn subgoals. Across Experiments 1 and 3, descriptions given by students in the label and visual-isolation groups of how to solve Poisson problems mentioned the goal to find the total frequency more often than did descriptions given by students in the no-label group. In Experiment 2, students in the label group preferred, more than did students in the no-label group, example solutions that separated the steps for finding total frequency from the step for finding λ. Students in the label group were also more likely than were students in the no-label group to circle the total frequency steps as a unit in their solutions to the test problems.

The superior transfer performance of the label and visual-isolation groups across the experiments is particularly impressive in light of the extensive elaboration provided to students in the study by Catrambone and Holyoak (1990). In that study, the overall transfer advantage for students receiving elaboration, although not reliable across all far transfer problems, was in line with performance by students in the label and visual-isolation groups in the present study who received a much more minimal manipulation. Thus, it appears that the label and visual isolation manipulations may be distillations of the information that provided the most benefit in the earlier study.

Related problem solving research by Reed et al. (1985) and Ross (1987, 1989) demonstrates that learners become attached to superficial details and mathematical procedures of examples in lieu of acquiring more generalized knowledge about how to solve problems in a particular domain. Although Reed et al. attempted to provide elaborations to help students go beyond the mathematical details of the training problems, they found poor performance on far transfer problems. The elaborations, though, may have failed to provide support for subgoal learning. The results of the present study suggest that a large amount of elaboration is certainly not the key to improving transfer from examples (see also Kieras & Bovair, 1984); rather, the additional information can be quite minimal if it focuses on the right kind of knowledge. This knowledge can be fruitfully conceptualized as subgoals.

Implications

The results strongly suggest that learners are often unable on their own to make a set of mathematical steps meaningful. Rather, students need help from either a teacher or textbook. This is an important finding because some educators may have a tendency to assume that the meaning of a set of steps is obvious and that the students will surely recognize their overall purpose. This assumption may be wrong much more often than educators would like to believe.

Even when educators recognize the value of subgoals, they might not be skilled at identifying the best ones for a particular problem-solving domain, especially in areas such as math and physics. For instance, Chi et al. (1989, p. 149) discussed a problematic mechanics example from a textbook. In the example, a block was suspended from a ceiling by two pieces of rope joined at a knot and by a third piece of rope going from the knot to the block. The task was to find the magnitude of two of the forces given the third force. The solution states that the knot where the three strings are joined should be considered the body. However, no explanation was given as to why this decision was made. The decision was made because to find a force in terms of other forces, the student must determine that the forces act on a common point. In this problem the only place where all three forces act was the knot. This critical subgoal of finding a common point where the forces were acting was information that would have been useful for students to have when solving future problems. However, instead of conveying this subgoal, the example was more likely to convey a series of steps that may or may not have been useful for other problems.

There may not exist a correct set of subgoals to be learned for a particular domain. Researchers might show one set to be more effective than another by looking at the problem-solving performance of students taught one set or the other. Thus, educators who differ in opinion about the usefulness of certain subgoals can explicitly compare their subgoal sets because the educators are at least using a common cognitive language.

The present study provides some information on this issue. Across the studies most students formed the subgoal to find λ. However, this subgoal was not all that useful to the students for solving novel transfer problems. If it were, then students in the no-label group should have successfully found λ in those problems as often as did the other students. That the no-label group did not suggests that learning the lower level subgoal of finding total frequency was crucial for transfer success. It is not clear whether an appropriate level can be proven or derived with a logical analysis or a cognitive architecture, but it is an interesting issue for future research.

Besides differing on which subgoals they believe are the most useful to teach, educators and researchers might also be unclear about the best ways to aid subgoal learning. The current results indicate that labeling and visual isolation can be effective techniques.

Extensions

Whereas the current results suggest that examples can be improved to aid problem-solving transfer, Chi and her colleagues (e.g., Chi & Bassok, 1989; Chi & VanLehn, 1991) suggested that educators can improve learning by teaching students to produce better self-explanations. Chi showed that individuals vary in terms of what kinds of information they extract from examples. In fact, students in the present study may have varied in what they learned from examples regardless of the experimental manipulations. That is, students who learn easily could have learned the goal to find the total frequency on their own when studying the examples regardless of the manipulation, whereas the students who have difficulty learning might have been less likely to do so (e.g., Chi et al., 1989). Unfortunately, no a priori information was collected by which to classify students in the experiments presented here. As a result, it is possible that the effects of the manipulations might have been clouded. In future experiments similar to the ones presented here, researchers could segregate students into those who learn easily and those who do not (by means of SAT scores or performance on some prior task) and examine whether the manipulations affect both groups in the same way.

Researchers could further test the subgoal approach by teaching learners subgoals that are hypothesized not to match well with the subgoals needed to solve test problems. These learners should perform as poorly as, or perhaps worse than, learners who simply memorized a set of steps. That is, if subgoals are used to guide problem solving performance, then inappropriate subgoals should hinder transfer.

It is suggested here that an important factor for educators interested in promoting transfer is to ensure that somehow the relevant subgoals for a domain are conveyed to learners. This implies that a crucial early step in teaching problem solving in a domain is for teachers to spend time identifying to themselves the useful subgoals from a novice's perspective. How might this be done? One possibility is to first identify a target set of problems that the instructor wants the students to be able to solve. Then the instructor should write out the solutions to these problems and analyze them to determine the subgoals achieved by groups of steps that constitute the solution procedures to the problems. Clearly, it will be important for researchers to find a standardized way of identifying subgoals that are to be taught. This might be a difficult task, but one that could greatly benefit teaching and learning.

References

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review, 86,* 124–140.

Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development, 1,* 103–121.

Catrambone, R. (1994a). Improving examples to improve transfer to novel problems. *Memory & Cognition, 22,* 606–615.

Catrambone, R. (1994b, January). *What in the heck are subgoals?* Paper presented at the Fifth Annual Winter Text Conference, Jackson Hole, WY.

Catrambone, R., & Holyoak, K. J. (1990). Learning subgoals and methods for solving probability problems. *Memory & Cognition, 18,* 593–603.

Chi, M. T. H., & Bassok, M. (1989). Learning from examples via self-explanations. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 251–282). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Bassok, M., Lewis, R., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182.

Chi, M. T. H., & VanLehn, K. (1991). The content of physics self-explanations. *Journal of the Learning Sciences, 1,* 69–106.

Dixon, P. (1987). The processing of organizational and component step information in written directions. *Journal of Memory and Language, 26,* 24–35.

Eylon, B., & Reif, F. (1984). Effects of knowledge organization on task performance. *Cognition and Instruction, 1,* 5–44.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 99–129). Hillsdale, NJ: Erlbaum.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8,* 255–273.

Lancaster, J. S., & Kolodner, J. L. (1988). *Problem solving and learning in a natural task domain* (Tech. Rep. No. GIT-ICS-89/02). Atlanta: Georgia Institute of Technology.

LeFevre, J., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction, 3,* 1–30.

Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs, 54*(Whole No. 248).

Mawer, R. F., & Sweller, J. (1982). Effects of subgoal density and location on learning during problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 252–259.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Pirolli, P. L., & Anderson, J. R. (1985). The role of learning from examples in the acquisition of recursive programming skill. *Canadian Journal of Psychology, 39,* 240–272.

Reed, S. K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 106–125.

Ross, B. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16,* 371–416.

Ross, B. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 629–639.

Ross, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 456–468.

Shaffer, J. P. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association, 81,* 826–831.

Smith, E. E., & Goodman, L. (1984). Understanding written instructions: The role of an explanatory schema. *Cognition and Instruction, 1,* 359–396.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 19–41). New York: Springer.

Ward, M. R., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7,* 1–39.

Zhu, X., & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4,* 137–166.

(Appendixes follow on next page)

Appendix A

Training Example Used in Experiment 1 With No-Label and Label Solutions

A judge noticed that some of the 219 lawyers at City Hall owned more than one briefcase. She counted the number of briefcases each lawyer owned and found that 180 of the lawyers owned exactly 1 briefcase, 17 owned 2 briefcases, 13 owned 3 briefcases, and 9 owned 4 briefcases. Use the Poisson distribution to determine the probability of a randomly chosen lawyer at City Hall owning exactly two briefcases.

a. No-Label Solution:

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X = x) = \frac{[(e^{-\lambda})(\lambda^x)]}{x!}$$

$$P(X = 2) = \frac{[(2.718^{-1.32})(1.32^2)]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

b. Label Solution:

Total number of briefcases owned = [1(180) + 2(17) + 3(13) + 4(9)] = 289

$$E(X) = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X = x) = \frac{[(e^{-\lambda})(\lambda^x)]}{x!}$$

$$P(X = 2) = \frac{[(2.718^{-1.32})(1.32^2)]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

Appendix B

Test Problem in Which Total Frequency Is Given Directly

A number of celebrities were asked how many commercials they made over the last year. The 20 celebrities made a total of 71 commercials. Use the Poisson distribution to determine the probability that a randomly chosen celebrity made exactly 5 commercials.

Solution (not seen by students):

$$E(X) = \frac{71}{20} = 3.55 = \lambda = \text{average number of commercials per celebrity}$$

$$P(X = 5) = \frac{[(2.718^{-3.55})(3.55^5)]}{5!} = \frac{(.029)(563.8)}{120} = .135$$

Appendix C

Test Problem in Which Total Frequency Is Calculated by Adding Simple Frequencies (Experiment 2)

Over the course of the summer, a group of 5 kids used to walk along the beach each day collecting seashells. We know that on Day 1 Joe found 4 shells, on Day 2 Sue found 2 shells, on Day 3 Mary found 5 shells, on Day 4 Roger found 3 shells, and on Day 5 Bill found 6 shells. Use the Poisson distribution to determine the probability of a randomly chosen kid finding 3 shells on a particular day.

Solution (not seen by students):

$$E(X) = \frac{4+2+5+3+6}{5} = \frac{20}{5} = 4.0 = \lambda = \text{average number of shells per kid}$$

$$P(X = 3) = \frac{[(2.718^{-4.0})(4.0^3)]}{3!} = \frac{(.018)(64)}{6} = .195$$

Appendix D

Test Problem With Weighted Average Method to Find λ but in Which Trials (Denominator) Must Be Found by Adding Number of Members in Each Category (Experiment 2)

A construction crew had a varying number of people who knew how to use a jackhammer, depending on the particular job that was needed. On 10 of the jobs they did, only one person knew how to use a jackhammer, on 13 of the jobs 2 people knew how to use jackhammers, on 6 of the jobs 3 people knew how to use jackhammers, and on 7 of the jobs 4 people knew how to use jackhammers. Use the Poisson distribution to determine the probability of exactly two people in the crew knowing how to use a jackhammer on a randomly chosen job.

Solution (not seen by students):

$$E(X) = \frac{1(10) + 2(13) + 3(6) + 4(7)}{10 + 13 + 6 + 7} = \frac{82}{36} = 2.28 = \lambda = \text{average number of “knowers” per job}$$

$$P(X = 2) = \frac{[(2.718^{-2.28})(2.28^2)]}{2!} = \frac{(.102)(5.2)}{2} = .265$$

Appendix E

Instance of Example Solution Studied by the Visual Isolation and Same Line Groups (Experiment 3)

a. Visual-Isolation Group

$$1(180) + 2(17) + 3(13) + 4(9) = 289$$

$$E(X) = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X = x) = \frac{[(e^{-\lambda})(\lambda^x)]}{x!}$$

$$P(X = 2) = \frac{[(2.718^{-1.32})(1.32^2)]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

b. Same-Line Label Group

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{\text{total number of briefcases owned}}{219} = \frac{289}{219}$$

$$= 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X = x) = \frac{[(e^{-\lambda})(\lambda^x)]}{x!}$$

$$P(X = 2) = \frac{[(2.718^{-1.32})(1.32^2)]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

Received February 22, 1994
Revision received July 5, 1994
Accepted July 18, 1994 ■

Generalizing Solution Procedures Learned from Examples[1]

Richard Catrambone
School of Psychology
Georgia Institute of Technology
Atlanta, Georgia 30332-0170
rc7@prism.gatech.edu

Abstract

Three experiments tested the hypothesis that when learners are led to group steps from example solutions, they will be more likely to learn subgoals that can be transferred to novel problems, thereby improving problem solving. The results from each experiment suggest that a label serves as a cue for grouping by demonstrating that a relatively meaningless label in an example solution was as effective (and sometimes more effective) as a meaningful label in helping learners transfer to novel problems. Experiment 2 provided converging evidence for subgoal learning by demonstrating that participants studying example solutions with a label were more likely to segment the solution as a function of the label and were more likely to mention the corresponding subgoal in their descriptions of how to solve problems.

Introduction

A good deal of research has examined the transfer success people have after studying training materials such as those containing step-by-step instructions (Kieras & Bovair, 1984; Smith & Goodman, 1984), examples (e.g., Ross, 1987, 1989), or both (Fong, Krantz, & Nisbett, 1986). Although there have been some exceptions (e.g., Fong et al., 1986; Zhu & Simon, 1987), the usual finding from such research is that people can carry out new procedures or solve new problems that are quite similar to those on which they were trained, but have difficulty when the novel cases involve more than minor changes from what they had previously studied.

This transfer difficulty seems to stem from a tendency by many learners to form representations of a solution procedure that consist of a linear series of steps rather than a more

[1]Currently under review.

structured hierarchy. An advantage of a hierarchical organization is that it can provide guidance for adapting the procedure for novel cases. One potentially useful hierarchical organization for a solution procedure would be a set of goals and subgoals with methods for achieving them (e.g., Anzai & Simon, 1979; Card, Moran, & Newell, 1983; Catrambone & Holyoak, 1990; Newell & Simon, 1972; Singley & Anderson, 1989). Problems within a domain typically share the same set of subgoals, although the steps for achieving the subgoals might vary from problem to problem. For instance, physics mechanics problems typically share the subgoals of identifying all "systems" in the problem and identifying all forces acting on the object of interest regardless of whether the problems involve objects on inclined planes or blocks suspended over pulleys (Heller & Reif, 1984).

Consider a student facing a novel problem, that is, one in which the steps are not the same as those seen in a previously-studied example. If the student has memorized only a rote set of steps for the overall solution procedure, he or she will have little guidance as to which steps need to be modified, as well as what new steps might need to be created, in order to solve the problem. Conversely, a student who learned a solution procedure organized by subgoals and methods could attempt to apply those subgoals to the novel problem. This approach has two advantages. First, the learner would know which steps from the learned procedure are relevant for achieving a particular subgoal. Thus, if those exact steps can not be carried out in the current problem, the learner knows which set of steps on which to focus for modification. Second, if the learner is attempting to achieve a particular subgoal and realizes that a modification to the old steps will not achieve the subgoal, then the subgoal can help constrain the memory search for other relevant information for achieving that subgoal (Anzai & Simon, 1979). Thus, the search space for useful information would be reduced for the second learner.

Forming Hierarchical Organizations for Solving Problems

A variety of studies have examined the effects on problem solving for learners who form a hierarchical structure representing a solution procedure (e.g., Dufresne, Gerace, Hardiman, & Mestre, 1992; Eylon & Reif, 1984). The typical result was that learners were able to solve novel problems more successfully than learners who were led to form a single-level organization of the problem solving procedure. In these studies, researchers usually derived what they believed to be a useful hierarchical approach to problem decomposition and induced learners to internalize this approach by following a prescribed method for solving or processing training problems. Although these hierarchical approaches were not always couched in terms of subgoals and methods, they could certainly be viewed in this light.

For instance, Heller and Reif (1984) formulated a model specifying the underlying knowledge and procedures needed to successfully solve mechanics problems. This procedural knowledge included drawing force diagrams for all external forces in a problem and identifying "short-range" and "long-range" forces. The authors required participants to solve three problems by adhering to a hierarchical model (the "M" model) for redescribing each problem in terms of relevant forces. This model was designed to address various deficiencies learners demonstrate when attempting to solve mechanics problems. The model was contrasted with the M* model that intentionally omitted certain levels of the redescription process hierarchy such as checking for consistency between the direction of forces and the resulting acceleration (learners were not prevented from doing these checks, but they were not reminded to do them). Participants who were required to follow the M model performed significantly better at redescribing novel problems and solving them.

Heller and Reif (1984) were quick to point out that the M model was only prescriptive and did not necessarily have any direct relationship to internal representations learners might have formed by following the model. It is also the case that besides containing certain high-level subgoals such as describing motion that the M* model lacked, the M model included a number of low-level "hints," such as reminding learners to include properties such as mass in their drawings, that the M* model did not include. Thus, it is not entirely clear whether requiring participants to focus on a certain prescribed set of subgoals was most responsible for superior performance on new problems or whether the lower-level procedural details were also crucial.

Brown, Kane, and Echols (1986) found a similar result in the area of short story problems. Children demonstrated better spontaneous transfer to novel problems if they had focused (either on their own or through an experimental intervention) on a hierarchical structure, in this case the goal structure, of the base analog.

Subgoal Learning

The studies mentioned above focusing on hierarchy learning used fairly powerful manipulations to induce learners to form different types of organizations. While this approach can yield practical implications for instructional manipulations, it provides fewer constraints for models of problem solving since it can be difficult to determine which features of the manipulation led to performance differences.

The purpose of the present research was to test a model of subgoal learning by examining the transfer effects due to a manipulation designed to influence whether learners formed a particular subgoal.

In the present work the assumption is made that learners' knowledge for solving problems in a domain can be fruitfully represented in terms of subgoals with steps for achieving them. At one extreme a learner could have a single goal: to solve a particular problem type. This goal would be achieved by a linear series of steps. Towards the other extreme would be a solution procedure broken into a number of subgoals, each with an associated method for achieving it. Each method would consist of a small number of steps relative to the total number of steps in the entire solution procedure.

Transfer to novel problems is assumed to be at least partly a function of applying subgoals learned from prior problems or examples to a new problem. Thus, manipulations designed to affect whether or not a particular subgoal is learned should affect a person's ability to solve novel problems.

The notion of a subgoal is sometimes applied to a construct generated by a learner when he or she reaches in impasse during problem solving (e.g., Newell, 1990, Chap 4; VanLehn, 1988) and is sometimes used to refer to a feature of a task structure that can be taught to a learner (e.g., Catrambone & Holyoak, 1990; Dixon, 1987). Both views converge on the prediction that learning subgoals can help one transfer more successfully to novel problems. The factors that affect subgoal learning remain unclear though.

<u>Factors Influencing Subgoal Learning</u>

Many models of transfer posit the existence of subgoals but do not necessarily explain how they are learned or formed. Anzai and Simon (1979) offered an account of subgoal learning in the context of a person learning to solve the Tower of Hanoi problem. They recorded the moves and verbal protocol of the learner as she solved the problem multiple times. Two observations from that study are related and particularly relevant to the present research. The first is that over trials the learner began to chunk groups of moves, that is, she would make a set of moves in quick succession followed by a pause before the next set of moves. The second is that she appeared to form goals and subgoals in her representation of the procedure for solving the Tower of Hanoi problem.

Anzai and Simon argued that subgoal acquisition is not trivial and its occurrence is greatly aided when the search space (e.g., possible moves in the Tower of Hanoi problem) is simplified. This simplification will frequently initially require prior knowledge of the learner of certain facts that can be applied to the domain. In the case of the Tower of Hanoi, such a fact might be that move repetitions are inefficient. When the search space is simplified, working memory load is reduced.

One hypothesized advantage of a working memory load reduction is that the learner is better able to notice and remember a sequence of steps that led to a particular outcome (see also Sweller, 1988). In Anzai and Simon's model this aids subgoal formation because a <u>sub</u>goal is formed when a learner is working towards a certain goal (perhaps derived from task instructions) and notices that a set of steps places him or her in a situation to be able to carry out additional steps that ultimately achieve the goal. The learner will be better able to notice the result of the first set of steps, and be able to chunk that sequence of steps, if working memory load has been reduced.

A key component in Anzai and Simon's model is the presence of a perceptual system that allows the learner to observe various external features of the problem situation. In the case of the Tower of Hanoi problem, one feature would be, for example, a particular disk being located next to a smaller disk. However, in learning tasks that are less obviously perceptually oriented, such as learning to solve word problems in probability, physics, or algebra, simple perceptual features are less likely to play a key role in subgoal formation. Rather, cues in worked examples and the learner's background knowledge will play a larger role. These cues may take the form of text and diagrams in the problem that direct the learner to relevant aspects of the problem and relevant prior knowledge (cf. Ward & Sweller, 1990). The cues can direct the learner to group a set of steps in the example solution and thus, to increase his or her chances of recognizing that a particular outcome is the result of the execution of those steps. That is, the recognition of the grouping is hypothesized to lead the learner to try to uncover the purpose of the group of steps. This "purpose" can be conceptualized as a subgoal. One such cue that could encourage grouping, and thus subgoal learning, is a label.

Labels Aid Categorization. Wattenmaker, Dewey, Murphy, and Medin (1986) found that providing learners with a theme (e.g., think of objects in one category as being or not being reasonable substitutes for a hammer) during a training session helped them learn categories more quickly (see also Cabrera & Billman, 1994; Homa & Cultice, 1984). Medin, Wattenmaker, and Hampson (1987) found that learners tended to sort a set of items around a single primary dimension or sometimes correlated features for which causal or explanatory links could be readily made. When themes (e.g., flying) were made more salient to learners, they tended to categorize more on family resemblance than individual features. These results suggest that features of example solutions that help learners form links, such as causal or explanatory ones, among items will help learners form a category that captures a useful relationship about the items. Thus, with respect to learning solution procedures, cues that help learners determine which steps go together and what their purpose is, might aid the formation of a subgoal representing the steps' purpose.

The hypothesized benefits of labels for subgoal learning are also consistent with Fried & Holyoak's (1984) category density model. Their model assumes that the goal of the category learner is to develop a schematic description of the distributions of category exemplars over a feature space. Highly salient features will tend to be encoded initially since, in an unfamiliar domain, the learner will probably have trouble identifying less salient features. In problem solving, these less-salient features constitute the deep structure of problems. For instance, when learners study examples in domains such as probability and physics, they tend to focus on familiar surface features such as basic mathematical operations and objects (e.g., multiplying a set of numbers, blocks on inclined planes) rather than goals being achieved by the operations such as determining the forces acting upon the object of interest (Chi, Feltovich, & Glaser, 1981). Manipulations that help learners focus on meaningful sets of steps could lead learners to discover the emergent deep structure.

How Subgoals Aid Transfer to Novel Problems

Wattenmaker et al. (1986) suggested that the degree of difficulty in learning a particular category structure is at least partly a function of the type of knowledge that learners bring to the task. With respect to the present work, subgoals influence the knowledge a learner brings to the task of solving a novel problem by reducing the search space that is explored while the learner tries to modify an old solution procedure.

A particular subgoal will have a set of steps, or method, associated with it, and if the learner discovers that those steps can not be used on a particular problem, he or she will have a reduced search space to consider when trying to adapt the method. That is, the learner knows on which steps to focus for changing the procedure. Conversely, a learner who has learned a solution procedure consisting of a single goal with a long series of steps will be less likely to determine successfully which steps need to be modified in order to solve the problem because the search space of possible modifications will be larger. In a more extreme case, if a particular subgoal needs to be achieved in a very different way than was demonstrated in the example (i.e., new steps are required rather than a modification of old steps), the learner possessing a representation with subgoals will have some guidance about what prior knowledge might be relevant to achieving the subgoal. The learner who memorized only a series of steps will be less likely to know what knowledge he or she possesses might be useful. Thus, a person representing a solution procedure in terms of subgoals and steps for achieving the subgoals has a more flexible procedure than a person representing the procedure as a single long series of steps.

For instance, in the probability examples used in the current study, the ultimate goal of each problem is to calculate a probability. The solution procedure for achieving this goal involves a number of steps, a subset of which constitutes a sequence of multiplication and addition operations that can be grouped under the subgoal "find the total frequency of the event."

Consider the "No Label" solution to the probability example in Table 1 involving the Poisson distribution.[1] A learner could study this example and memorize the steps for solving a problem that involves the same set of steps even if the new problem involved farmers and tractors instead of lawyers and briefcases. After studying the No Label solution, the learner's knowledge for the part of the solution procedure that involves finding λ, the average, might be represented as:

<u>Goal: Find λ</u>

Method: 1. Multiply each category (e.g., owning exactly zero briefcases, owning exactly one briefcase, etc) by its observed frequency.

2. Sum the results.

3. Divide the sum by the total number of lawyers to obtain the average number of briefcases per lawyer.

This representation would serve the learner well for problems that involve calculating λ in the same way as the example. However, this representation fails to capture the fact that the first line of the No Label solution in Table 1a also involves calculating a total frequency. Finding the total frequency is a subgoal that might be achieved in a variety of ways depending on the givens in the problem. A novel problem that requires finding total frequency in a different way than in the example might cause problems for the learner with the above representation. For instance, consider the problem in Table 2a. In this problem the total frequency is calculated by adding a set of simple frequencies. This is a less-complex method than was used in the example, but the learner might not be able to construct it because the subgoal for finding the total frequency, and an instance of a method for achieving it, were never isolated. If the learner had formed the following representation, then his or her chance of solving the problem in Table 2a might be better since this representation identifies the steps involved in finding the total:

<u>Goal: Find λ</u>

Method: 1. <u>Goal: Find total number of briefcases</u>

Method: a. Multiply each category by its observed frequency.

b. Sum the results to obtain the total number of briefcases.

2. Divide the total number of briefcases by the total number of lawyers to obtain the average number of briefcases per lawyer.

Insert Tables 1 and 2 about here

Catrambone (in press) found that learners studying the "Meaningful Label" solution (Table 1b) to the example were more likely than No Label learners (who studied the solution in Table 1a) to find the total frequency as measured by their success at solving problems such as the one in Table 2a. This was taken as initial evidence that the former group had learned the subgoal to find the total frequency.

The account proposed in Catrambone (in press) for why the Label group was more likely than the No Label group to learn the subgoal to find the total frequency could be summarized as follows:

1) A label leads learners to group a set of steps (such as the steps for finding the total frequency).
2) After grouping the steps, learners are likely to try to self-explain why those steps go together.
3) The result of the self-explanation process is the formation of the goal that represents the purpose of that set of steps.

While most learners can presumably engage in a self-explanation process, with varying degrees of success, good students seem better at determining the appropriate boundaries between meaningful groups of steps in a solution procedure (Chi, Bassok, Lewis, Reimann, & Glaser, 1989). The use of a label in examples is hypothesized to serve as a cue to the boundaries. Thus, Label participants in Catrambone (in press) were helped presumably to focus on the steps that formed a coherent unit (finding a total) whereas the No Label participants were not.

While the results of Catrambone (in press) were consistent with the above account, they did not constitute a strong test of the account. That is, it is possible that part of the transfer advantage enjoyed by the Label group could have been due to the fact that the label itself provided information beyond serving as a cue to group a set of steps. That is, the label indicated that the total number of briefcases was being found. Thus, instead of the label leading learners to group a set of steps and inducing a self-explanation process, it may simply have provided them with this fact: finding the total number of things is something that one does when solving Poisson problems (it is assumed that college-aged learners are sophisticated enough to generalize "briefcase" to "things").

One way to tease apart these possible explanations is to provide learners with labels that contain no explicit information about the domain and examine whether transfer performance is as good as transfer performance by learners who study examples with more meaningful labels. This is the general approach taken in the experiments in the present study.

Experiment 1 explored the possibility that a label has to have some meaning in order for the learner to learn a subgoal as opposed to a simply serving as a cue for the learner who then uses background knowledge to construct the subgoal. Experiment 2 sought converging evidence for subgoal learning to supplement the evidence from transfer performance. Experiment 3 explored the generality of the subgoal that is formed from a semantically meaningful label compared to one formed from a less-meaningful label.

Learning from Examples

The present study examined learners acquiring subgoals by studying examples rather than through direct instruction of those subgoals. There were two primary reasons for this approach. First, people typically prefer to study and use worked examples, or problems they have previously solved, when attempting to solve new problems (e.g., LeFevre & Dixon, 1986; Pirolli & Anderson, 1985) despite the fact that they are frequently misled by surface features of the examples and problems (e.g., Ross, 1987, 1989).

Second, there is evidence that students, at least good students, can derive knowledge from examples that they did not or could not acquire from explanatory text, even if the text was formally complete. Chi et al. (1989) found that after studying a text on mechanics, good and poor students (as defined by a subsequent problem solving test) seemed to possess similar declarative knowledge. However, after studying worked examples, good students were more likely to acquire knowledge about, among other things, the conditions of application of actions or operators, the consequences of these actions, and the relationship of the actions to goals. Chi et al. suggested that this additional problem solving or procedural knowledge was the result of a self-explanation process. Good students were far more likely than poor students to produce self-explanations leading to the acquisition of this knowledge. One reason good students were more likely to do this is because they could better recognize the locations in examples that contained unexplicated actions.

From an educational point of view, there are two responses to the findings of Chi et al. One is to find ways to help poor students improve their self-explanation skills so that they can derive more from examples. The second, and the one chosen here, is to find ways to improve examples to help all learners derive useful information, such as subgoals, from examples. In the present work, this is done through cues in example solutions that could help learners group steps that they otherwise might not realize were related.

Experiment 1

Experiment 1 examined the role of label meaningfulness in subgoal learning. In this experiment the No Label group studied examples demonstrating the weighted average method for finding λ (all solutions in Table 1 demonstrate this type of solution). The Meaningful Label and Less-Meaningful Label groups' examples differed in that the steps for finding the total frequency were explicitly labeled rather than merged with the overall set of steps for finding λ (see Tables 1b and 1c for instances of the "Meaningful Label" and "Less-Meaningful Label" solutions, respectively).

Subgoal learning was assessed by how successfully participants found λ on test problems such as those in Table 2. The presence of a label was not hypothesized to improve learners' ability to memorize the solution procedure since they were allowed to work at their own pace. Thus, the No Label group was predicted to be as successful as the label groups at solving transfer problems that were isomorphic to the training examples. If learners receiving examples with labels solve novel problems more successfully than the No Label group, this would be consistent with the hypothesis that the presence of a label, rather than its meaningfulness, was sufficient to induce subgoal learning. If the Meaningful Label group outperforms the Less-Meaningful Label group, this would suggest that the meaningfulness of the label affected the likelihood of forming a domain-relevant subgoal or aided transfer in some other way.

Method

Participants. Participants were 100 students recruited from an introductory psychology class at the Georgia Institute of Technology who received course credit for their participation. None of the students had taken a probability course prior to participating in the experiment.

Materials and Procedure. All participants initially studied a cover sheet that briefly described the Poisson distribution and how it could be used as a replacement for more cumbersome techniques for calculating probabilities involving events that could be categorized as successes and failures. The Poisson equation was presented along with a simplified notion of a random variable.

Participants were randomly assigned to one of three groups. The Meaningful Label group (N = 34) studied three examples demonstrating the weighted average method for finding λ in which the steps for finding the total frequency were given a label that was assumed to have meaning to the participants and made mathematical sense given the steps that preceded it (see the "Meaningful Label Solution" in Table 1b for an example). The Less-Meaningful Label group (N = 34) studied examples in which the steps for finding the total frequency were labeled with Ω which was assumed to have little meaning for the participants in the context of the examples (see the "Less-Meaningful Label Solution" in Table 1c). The No Label group (N = 32) studied examples in which the steps for finding the total frequency were not labeled (see the "No Label Solution" in Table 1a).

After studying the examples, participants solved six test problems. The first two required the use of the weighted average method for finding λ (isomorphic to the example in Table 1). These problems were given first so that participants would be able to immediately see that the prior examples were relevant for solving the test problems. The third and fourth problems provided the total frequency directly, thus λ could be found by simply dividing the given total frequency by the total number of trials. These problems were the problem in Table 2b and another problem isomorphic to it. The fifth and sixth problems involved adding simple frequencies in order to find the total frequency (see Table 2a for one of the problems). Participants were told not to look back at the examples when solving the test problems.

Participants' written solutions were scored for whether they found λ correctly.

Design. The independent variable was type of example solution studied (Meaningful Label, Less-Meaningful Label, No Label), thus there were three groups in the experiment. The dependent measure was performance on the six test problems, two of which were isomorphic to the examples while the other four were novel.

Results

Participants were given a score of 1 for a given problem if they found λ correctly and a score of 0 otherwise. The scores for the two problems that were isomorphic to the training examples, Problems 1-2, were summed, thus creating a score from 0-2 for performance on those problems. Similarly, the scores for the four novel problems, Problems 3-6, were summed, thus creating a score from 0-4 for performance on those problems.

As expected, all groups did quite well at finding λ on test problems that were isomorphs to the training examples (Problems 1-2). All participants except two, one in the Less-Meaningful Label group and one in the No Label group, solved both isomorphs correctly.

There were significant differences among the three groups with respect to finding λ in the novel test problems, $F(2, 97) = 10.20$, $p = .0001$, $MS_e = 1.82$, with means of 3.44, 2.97, and 1.97 for the Meaningful Label, Less-Meaningful Label, and No Label groups, respectively. The most typical mistake that students made on these problems was to write in the solution area that not enough information was given to solve the problem. Shaffer (1986) sequential Bonferroni pairwise comparisons (familywise $\alpha = .05$) indicated that both label groups outperformed the No Label group, both ps $< .05$. No reliable performance difference was found between the two label groups ($p = .13$).

Discussion

The transfer performance of the three training groups reveals that the presence of a label facilitates performance on novel problems. These results are consistent with the hypothesis that a label helps learners form a subgoal for the labeled steps. Even a relatively meaningless label appears to promote subgoal learning, presumably because it serves as a cue to learners to retrieve information from long-term memory in order to explain why a set of steps belong together.

It is proposed that a label aids transfer because it leads learners to group a set of steps and then to self-explain why the steps go together. The result of this self-explanation process is the subgoal. The results from Experiment 1, while demonstrating the relationship between labels and improved transfer, did not directly test the hypothesized links of the model. Experiment 2 began a more direct exploration of the links between labeling, grouping, and subgoal formation.

Experiment 2

Experiment 2 examined the hypothetical connections among labeling, grouping, and subgoal formation, as well as continuing to examine the relationship between labeling and transfer performance, through the use of three tasks. The first task was a segmenting task in which participants circled the steps in a worked-example that they believed went together. It was hypothesized that participants whose examples provided a label for the set of steps for finding the total frequency would be more likely to circle those steps as a unit, that is, to group those steps, compared to the No Label participants who might be more likely to circle the entire set of steps for finding λ as a unit.

The second task required participants to provide a description of how to solve problems in the domain. It was hypothesized that if a person learns a subgoal, such as the subgoal to find a total, then he or she would be more likely to mention it in a description compared to a learner who did not learn that subgoal. That is, the description would be more hierarchically organized compared to one that simply listed steps.[2]

The third task was transfer performance on novel problems.

Thus, participants' segmentation performance, descriptions, and transfer performance were used as converging measures of subgoal learning.

Method

Participants. Participants were 90 students recruited from an introductory psychology class at the Georgia Institute of Technology who received course credit for their participation. None of the students had taken a probability course prior to participating in the experiment.

Materials and Procedure. Participants studied the same cover sheet as in Experiment 1 and were randomly assigned to one of three groups. The Meaningful Label ($N = 30$) and Less-Meaningful Label ($N = 30$) groups studied the same three examples as the corresponding groups in Experiment 1 (in which the steps for finding the total frequency were explicitly labeled) while the No Label group ($N = 30$) studied the same examples used by the No Label group in Experiment 1.

The solutions studied by the label groups were modified from Experiment 1's so that the steps for finding the total frequency were on the same line as the rest of the steps for finding λ, thus making the solutions more visually similar to those studied by the No Label group (see Table 3 for an example). There were two reasons for the modification. The first was to reduce the chance that any segmenting and transfer differences could be due to factors other than labeling. The second was to guard against the possibility that participants in the label groups would circle a series of steps simply because they were on their own line.

Insert Table 3 about here

Accompanying the third example were instructions asking participants to circle steps in the solution that they felt formed a unit. The actual instructions were:

> In the solution presented below, please circle the groups of steps that you feel go together. For instance, suppose you were following a recipe for cooking something. Perhaps the first three steps of the recipe involved putting various ingredients into a bowl and the fourth step involved stirring the ingredients with a spoon and the fifth step involved using a blender to finishing the mixing. You might draw a circle around the first three steps because they involve "adding ingredients" and you might draw a circle around the the fourth and fifth steps because they involve "blending the ingredients."

After performing the segmenting task, participants were asked to describe how to solve problems in the domain. The instructions were:

Suppose you were going to teach someone how to solve Poisson distribution problems of the types you have just studied. Please describe the procedure or procedures you would give someone to solve these problems. Please be as complete as possible.

After writing their descriptions, participants solved the same six test problems used in Experiment 1. Participants were told not to look back at the examples when writing their descriptions or solving the test problems.

Participants' segmenting performance was scored for whether they circled the steps for finding the total frequency as a unit. Participants' explanations of how to solve problems in the domain were scored for two features: an explicit mention of trying to find the total and an explicit mention of trying to find an average. Participants' written solutions to the test problems were scored for whether they found λ correctly.

Two raters independently scored the explanations and agreed on scoring 94% of the time. Disagreements were resolved by discussion.

Design. The independent variable was type of example solutions studied (Meaningful Label, Less-Meaningful Label, No Label), thus there were three groups in the experiment. The dependent measures were segmenting performance, descriptions for how to solve the problems, and transfer performance on the six test problems.

Predictions

Segmenting. If the label manipulation made learners more likely to group a set of steps, then the label groups should tend to circle the steps (in the third example) for finding total frequency as a single unit more often than No Label participants.

Description of How to Solve Problems. Since it was hypothesized that the label groups would be more likely than the No Label group to learn the subgoal of finding a total, the label groups should be more likely to mention the idea of finding a total in their descriptions of how to solve Poisson problems. The Less-Meaningful Label group might mention the notion of finding

"Ω" rather than finding the "total number" of things if there is some tendency by learners to repeat the wording from examples. All groups were expected to mention finding λ equally often since this subgoal was always labeled in the examples.

Transfer. As in Experiment 1, the label groups were predicted to be more likely to find λ successfully on novel test problems since they involved new ways of finding the total frequency. All groups were predicted to find λ correctly in the isomorphic test problems since the same sets of steps used in the examples could be applied to those problems.

Results

Segmenting. As predicted, the label groups circled the steps for finding total frequency as a single unit more frequently than the No Label group, $\chi^2(2) = 6.57$, $p = .037$ (see Table 4).

Descriptions of How to Solve Problems. As expected, there was no significant difference among the groups in the frequency of mentioning the subgoal of finding λ or the average, $\chi^2(2) = 0.69$, $p = .71$ (see Table 4).

The frequency with which participants mentioned the subgoal of finding a total was analyzed in two ways. In both analyses, participants in the Less-Meaningful Label group who mentioned the notion of finding Ω and also explicitly called this value a total were counted in Table 4 as having mentioned a total rather than Ω.

The first analysis used the Total vs. Ω vs. Neither breakdown in Table 4 and found a significant difference among the groups, $\chi^2(4) = 28.14$, $p < .0001$. The percentages in Table 4 suggest that the label groups were more likely to mention the subgoal of finding a total or Ω compared to the No Label group. In the second analysis the Less-Meaningful Label participants who mentioned finding Ω but did not also call it a total ($N = 8$) were placed in the same category as those who explicitly mentioned a total. Thus, all participants were categorized into one of two groups: those mentioning a total or Ω versus those mentioning neither. Once again there was a significant difference between the groups in the frequency of mentioning total / Ω, $\chi^2(2) = 14.52$, $p = 0007$.

Insert Table 4 about here

Transfer. Test problems were scored as in Experiment 1.

All participants except one in the Meaningful Label group solved both isomorphic problems correctly.

As predicted, there were significant differences among the three groups with respect to finding λ in the four novel test problems, $F(2, 87) = 14.93$, $p = .017$, $MS_e = 3.50$, with means of 3.07, 2.80, and 1.73 for the Meaningful Label, Less-Meaningful Label, and No Label groups, respectively. Shaffer (1986) sequential Bonferroni pairwise comparisons indicated that both label groups outperformed the No Label group, both $ps < .05$. No reliable performance difference was found between the two label groups ($p = .57$).

Transfer to Novel Problems as a Function of Segmenting Performance. Participants who circled the steps for finding total frequency as a single unit on the third training example ($N = 50$) outperformed those who did not circle those steps as a unit ($N = 40$) with respect to performance on the novel test problems, $F(1, 88) = 62.72$, $p < .0001$, $MS_e = 3.09$, $Ms = 3.28$ and 1.60, respectively.

Transfer as a Function of Descriptions. There was a significant difference in transfer performance for novel problems as a function of whether participants mentioned finding a total ($N = 35$), Ω ($N = 8$), or neither ($N = 47$) in their descriptions (Total = 3.31, $\Omega = 3.50$, Neither = 1.79), $F(2, 87) = 27.49$, $p = .0004$, $MS_e = 3.21$. Pairwise comparisons indicated that participants mentioning a total or Ω did not differ from each other in transfer performance, and that both groups outperformed those who did not mention either, both p's < .05 (although the small number of participants mentioning only Ω makes interpretations of comparisons with them tentative).

Participants who mentioned finding the average ($N = 59$) did not perform differently from those who did not mention finding it ($N = 31$), $F(1, 88) = 0.56$, $p = .46$, $MS_e = 3.75$, $Ms = 2.64$ and 2.32, respectively.

Discussion

The purpose of Experiment 2 was to determine whether more direct support could be found for the hypothesized connections among labeling, grouping, and subgoal formation. The segmenting results provide initial support for the link between labeling and grouping, and the description results support the connection between labeling and subgoal formation. Finally, the transfer results are consistent with the claim that subgoal formation aids transfer. In sum, converging evidence was found for a connection between labeling and subgoal formation.

While results from the first two experiments are consistent with the "label as grouping cue" view, the generality of the subgoal formed as a function of the label is unclear. That is, a subgoal that is formed in response to a label that makes mention of superficial features in the example might become tied to those features. For instance, the subgoal formed by Meaningful Label participants in the first two experiments might have been "find the total number of objects." Conversely, a subgoal formed in response to a more abstract label might be less likely to be tied to superficial features. For instance, the subgoal formed by Less-Meaningful Label participants might have been "find the total." This latter subgoal is more general and closer to being formally correct.[3] Most participants in Experiment 2 who mentioned finding a total in their descriptions usually mentioned the total in terms of objects or things. However, it is possible that this surface feature tie was strongest for participants in the Meaningful Label condition.

One implication of forming a subgoal that is tied to superficial features is that the learner is confusing superficial and structural features of the domain. A way to test this possibility is to construct test problems that systematically manipulate the relationship between superficial and structural features and observe the degree to which the features guide learners' performance.

For instance, Ross (1987, 1989) provided students with various types of probability examples to study such as problems dealing with permutations and combinations. The permutation examples involved people picking objects in a certain order. Because the problems involved people picking objects, the number of objects in the problem provided the starting value for the denominator in the permutation equation. Some of the test problems involved people being assigned to objects. In these cases the number of people in the problem provided the starting value for the denominator. However, students typically placed the number of objects in the denominator. That is, students in Ross' studies appeared to confuse the superficial features of humans and objects with the domain-relevant features of choosing and chosen.

With respect to the experimental materials used in the present study, most learners, at least at the college level, were assumed to be sufficiently sophisticated to generalize "total number of briefcases." The generalization that might be formed though was unclear. One possibility was that the generalization would be "total number of objects" if the examples involved humans using objects. Learners forming this generalization would be predicted to be more successful solving novel problems that require the total number of objects to be calculated in new ways compared to learners not forming this generalization. However, given that this generalization is still tied to a superficial feature, objects, these learners might fail to solve correctly a novel problem that required the number of humans rather than objects for the total. Learners studying examples with the less-meaningful label who form the subgoal for finding a total might be less likely to have this subgoal tied to a superficial feature. As a result, these learners would be less likely to make mistakes on novel problems that switch the roles of humans and objects from their roles in the training examples.

This possibility was explored in Experiment 3.

Experiment 3

In Experiment 3 participants were divided into the same groups as the first two experiments. However, additional test problems were created in which the total frequency was calculated using people rather than objects.

Performance predictions varied as a function of training condition and type of test problem. The first four test problems were isomorphic to the training examples. The first pair involved calculating the total number of objects in order to find λ and the second pair involved calculating the number of people. Participants were expected to solve the first pair with little difficulty.

Participants were also expected to solve the second pair successfully even though they involved a reversal in the roles of humans and objects. The reason involves two types of superficial similarity that seem most likely to affect performance. The first is the format of the numbers in the examples and problems. For instance, consider the problem in Table 5a in which the roles for humans and objects are reversed from the examples (such as the one in Table 1). In both cases there are a series of numbers (one through four for the example in Table 1 and one through five in the problem in Table 5a) in which each number has another value associated with it. Each number is multiplied by its associated value and the products are added. This format similarity allows a number-matching approach that has been shown to be a powerful factor in transfer performance (Novick & Holyoak, 1991). In the present case, this number-matching approach would produce the correct answer. The second type of superficial similarity is the roles of humans and objects. They are reversed, compared to the examples, in the second pair of isomorphs. This reversal could, in some circumstances, lead learners to place incorrect values into an equation.

It was expected that the first type of similarity would drive performance because the isomorphic problems do not supply obvious candidates to place in the equation other than those produced via number-matching. Thus, regardless of whether a solver noticed or understood the role-reversal, he or she would still be likely to follow a number-matching approach and thus, produce the correct answer. For instance, for the problem in Table 5a, the learner could still follow the training procedure of multiplying "1" by the number nearest it, "2" by the number nearest it, and so on. This approach would also be consistent with prior observations that learners tend to prefer to use learned sequences of steps when they are allowable in new problems even if they are not optimal (cf. Luchins, 1942; Singley & Anderson, 1989, p 99).

The above approach would <u>not</u> work for the transfer problems that involve a change in steps such as those in Tables 2 and 5b and 5c. There is no way to carry out the old set of steps to solve these problems. It is assumed that in the process of trying to decide how to solve these problems, the learner will be more likely to notice features of the problem including which numbers are associated with humans and which are associated with objects. As a result, the solver with a subgoal linked to superficial features might be more likely to want to find or calculate the total based on objects even for the reversed-correspondence problems.

Based on the above analysis, predicted performance on the novel problems varied as a function of group and role-correspondence. Four of the novel problems provided the total frequency directly (such as the problem in Table 2b). Two of the problems involved objects providing the total event frequency and two involved humans providing the total event frequency (see Table 5b for an example of the latter type). It was predicted that on the first pair of problems both label groups would outperform the No Label group. It was predicted that on the second pair of problems Meaningful Label participants would be more likely than Less-Meaningful Label participants to incorrectly place the number of objects in the numerator of the fraction in order to find λ. This prediction was made because of the hypothesis that Meaningful Label participants would be more likely than Less-Meaningful Label participants to associate objects with finding a total. The value placed in the numerator of the fraction presumably represents the value that the participant believes is the total.

The last four problems involved adding simple frequencies in order to find a total frequency (such as the problem in Table 2a). Two of the problems involved objects being used to calculate the total event frequency and two involved humans being used to calculate the total event frequency (see Table 5c for an example of the latter type). It was predicted that on the first pair of problems the label groups would outperform the No Label group. The second pair of problems provided two sets of numbers. The first set of numbers could be added to produce a total number of objects and the second set could be added to produce a total number of people. If only one set was provided, then participants would be more procedurally constrained (as with the isomorphs) and there would be less of a chance of finding a performance difference between Meaningful and Less-Meaningful Label participants. It was predicted that on the second pair of problems the Meaningful Label participants would be more likely than Less-Meaningful Label participants to make the mistake of calculating a total using objects rather than humans.

Thus, for "reversed correspondence" novel test problems, it was predicted that Less-Meaningful Label participants would show less of a decrement in performance than Meaningful Label participants, relative to the groups' performance on the "same-correspondence" problems.

Insert Table 5 about here

Method

Participants. Participants were 90 students recruited from an introductory psychology class at the Georgia Institute of Technology who received course credit for their participation. None of the participants had taken a probability course prior to participating in the experiment.

Materials and Procedure. The conditions, training procedure, and cover sheet were identical to those in the prior experiments and there were 30 participants in each condition. The examples used the format created in Experiment 2 that made the visual appearance of the solutions used in the label conditions more similar to the No Label solution.

A larger number of test problems (12) were used than in the previous experiments. The first four test problems were isomorphic to the training examples. The first two involved objects in the total frequency and the next two involved humans in the total frequency (see Table 5a for an example of the latter set). The next four test problems involved the total frequency being given directly in the problem. The first pair of the set involved objects in the total frequency (see Table 2b for an example) while the second pair involved humans in the total frequency (see Table 5b for an example). The next four test problems involved calculating the total frequency by adding a set of simple frequencies. The first pair of the set involved objects in the total frequency (see Table 2a for an example) while the second pair involved humans in the total frequency (see Table 5c for an example).

Participants' written solutions were scored for whether they found λ correctly.

Design. The between-subjects variable was type of example solutions studied (Meaningful Label, Less-Meaningful Label, No Label) and the within-subjects variable was correspondence of the roles of humans and objects in the test problems to their roles in the examples. The dependent measure was performance on the 12 test problems.

Results and Discussion

As in the prior experiments, participants were given a score of 1 for a given problem if they found λ correctly and a score of 0 otherwise. The scores for Problems 1 and 2, the two problems that were isomorphic to the training examples and had the same role-correspondence of humans and objects as the examples, were summed, thus creating a score from 0-2 for performance on those problems. Similarly, a score from 0-2 was calculated for the isomorphs that had a reversed role-correspondence of humans and objects (Problems 3 & 4). Finally, a score from 0-4 was calculated for the novel test problems with the same role-correspondence as the examples (Problems 5, 6, 9, & 10) and a score from 0-4 was calculated for the novel test problems with a reversed role-correspondence (Problems 7, 8, 11, & 12).

As expected, all groups did quite well at finding λ on test problems that were isomorphs to the training examples regardless of whether the roles of humans and objects were reversed from the examples. In fact, all participants solved each of these problems correctly.

Table 6 presents the groups' performance on the novel test problems as a function of whether the problems involved the same or reversed role-correspondence of humans and objects compared to the training examples. An analysis of variance was carried out on the performance on the novel test problems with group as the between-subjects variable and role-correspondence (same as examples vs. reversed from the examples) as the within-subject variable.

Consistent with the results from the prior experiments, there was a significant difference among the three groups with respect to finding λ on the novel test problems, $F(2, 87) = 6.21$, $p =$

.003, $MS_e = 5.70$. There was also an effect of role-correspondence indicating that problems with reversed role-correspondence were solved less successfully than those with the same role-correspondence as the examples, $F(1, 87) = 26.56$, $p < .0001$, $MS_e = 0.86$. Finally, there was a significant interaction between group and role-correspondence, $F(2, 87) = 6.25$, $p = .003$, $MS_e = 0.86$, suggesting that the correspondence manipulation affected the groups differently.

Separate analyses were carried out for each group comparing performance on same and reversed role-correspondence problems. The Shaffer (1986) procedure for providing a familywise α of .05 for multiple comparisons was used. The Meaningful Label group showed a significant decrease in performance on the reversed role-correspondence problems compared to the same role-correspondence problems, $F(1, 29) = 19.12$, $p = .0001$, $MS_e = 1.54$, while the Less-Meaningful Label and No Label groups did not show significant differences in performance on the problem types (ps > .05).

Consistent with the above analysis, if performance on only the reversed role-correspondence problems is considered, a significant effect of group is found, $F(2,87) = 5.33$, $p = .007$, $MS_e = 3.41$, with pairwise comparisons indicating that the Less-Meaningful Label group outperformed the other groups (both ps < .04) but the Meaningful Label group did not outperform the No Label group ($p > .26$).

It was predicted that a typical mistake made by the Meaningful Label participants in solving the reversed role-correspondence problems would be to put or calculate a value for total number of objects in the numerator. One way of examining the likelihood of making this mistake is to examine performance on reversed role-correspondence problems by participants who solved the same role-correspondence problems correctly. This approach would therefore consider only participants who demonstrated the ability to transfer to problems that involved a change in procedure relative to the training examples. Thus, mistakes on the reversed role-correspondence problems would presumably be due to mapping problems rather than other sorts of transfer difficulties.

Of the eight Meaningful Label participants who found λ correctly in all four same role-correspondence problems, five of them put objects in the numerator for finding λ in the reversed role-correspondence problems. Conversely, of the seven Less-Meaningful Label participants who found λ correctly in all four same role-correspondence problems, only two of them put objects in the numerator for finding for finding λ in the reversed role-correspondence problems. While these numbers are too small to achieve statistical significance, their pattern is consistent with the interpretation that the Meaningful Label participants who were able to adapt the solution procedure from the examples were more likely to be misled by superficial features compared to the analogous Less-Meaningful Label participants.

Insert Table 6 about here

It is possible that the results from this experiment, rather than reflecting learners forming a subgoal associated with a superficial features, are due to learners including superficial features in the steps for finding the total. The present results do not allow these possibilities to be clearly disentangled. The data most relevant to examining this issue, the descriptions provided by learners, tended to contain a mention of objects at both the step level and at the subgoal level. Nevertheless, the transfer results suggest that participants receiving meaningful labels associated superficial features and mathematical roles more often than participants receiving less-meaningful labels.

General Discussion

Students frequently learn a solution procedure as a series of steps with little or no higher-level organization (Reed, Dempster, & Ettinger., 1985). As a result, while they can solve new problems that involve the same steps as a previously-studied example, they have difficulty with problems that require a change in the steps, even though the conceptual structure from the example to the problem is preserved.

A guiding assumption of the present research is that transfer performance will be enhanced if a solution procedure is structured by subgoals and a method for achieving each one rather than just a single linear set of steps for the entire procedure. Presumably there is a continuum of structuredness depending on the number of subgoals into which a procedure is broken.

A solution strategy containing a single goal and a set of steps that occur in a predetermined order to achieve that goal is similar to a rote strategy. Such a strategy is not particularly flexible in terms of supporting transfer to novel problems. Nevertheless, Singley and Anderson (1989) point out that there are advantages to a rote procedure: it is more efficient in terms of number of rule firings (if one were to model the procedure with production rules) and a rote procedure is often easier to learn and perform than a procedure that has more structure to it. Thus, it seems plausible that learners will tend to form a rote procedure unless induced to do otherwise. Factors that may lead learners to form more structured or hierarchical representations of procedures include their ability to articulate goals that sets of steps are achieving (Chi et al., 1989), the learner's reasoning style (e.g, Dufresne et al., 1992), and the type of training materials used (e.g., Catrambone, 1994; Eylon & Reif, 1984).

Based on prior work involving instructional manipulations (e.g., Eylon & Reif, 1984; Smith & Goodman, 1984), problem solving (Anzai & Simon, 1979), and categorization (e.g., Wattenmaker et al., 1986), I hypothesized that learners would be more likely to learn a subgoal from an example if the steps for achieving that subgoal were labeled. A label was predicted to make a learner more likely to group the set of steps and, through a self-explanation process, form a subgoal that represented the purpose of the steps. The focus of the experiments was primarily on the connection between labeling, subgoal formation, and transfer performance. The self-explanation process was not directly examined.

Experiments 1 and 2 demonstrated that the presence of a label, rather than its semantic content, can be sufficient to induce a learner to form a subgoal. Participants with less-meaningful labels were able to solve transfer problems as successfully as participants with meaningful labels and both groups transferred better than a No Label group. In addition, the segmentation task results were consistent with the claim that a label can aid grouping. These results, coupled with the description data and transfer performance, support the hypothesized links between labeling, grouping, and subgoal formation.

Experiment 3 demonstrated that the subgoal formed in response to a label may be less tied to superficial details of examples when the label does not reference those details. Learners receiving examples using less-meaningful labels transferred more successfully than other learners to novel problems that altered, with respect to the examples, the role-correspondence between superficial features and the solution procedure. This result, besides again supporting the view that labels serve as a grouping cue, suggests that the generality of the procedure formed from examples can be increased through the use of labels that do not contain references to superficial features of examples.

A related explanation for the obtained results is that the Less-Meaningful Label participants had to discover the purpose of the labeled steps as opposed to being told. The strategies or processes involved in determining the purpose of the steps might be related to the processes for constructing new steps or modifying old steps to achieve the same purpose or subgoal (e.g., McDaniel & Schlager, 1990).

Instructional and Individual Differences Factors Affecting Subgoal Learning

To be sure, the background of a learner plays a role in how likely he or she is to learn a subgoal. Ausubel (1968, p 148-149) suggested that the value of "organizers" hinges upon the learner possessing relevant background information so that the pieces of information being organized already have some meaning. For instance, if a student learning mechanics is told that one part of a solution procedure is to determine the components of force along the x and y axes, this organizer for the subsequent steps will be of minimal use if the learner knows little or nothing about coordinate systems or trigonometry.

With respect to the present study, a learner with a weak math background may look at a series of addition and multiplication steps labeled with Ω and not group them or, in grouping them, not realize that the steps calculate a total. This learner might be predicted to be less likely to form the subgoal of finding the total number of objects in this situation compared to one in which the steps were labeled with "total number of briefcases owned." For this learner, meaningful labels might produce better subgoal learning since the extra domain information provided by a meaningful label could help the learner make sense of the steps and to understand their purpose (even though the resulting subgoal might have ties to superficial features). Conversely, a learner with a stronger math background might be expected to recognize that the series of steps, when separated from the other steps in the overall solution procedure through the use of a label such as Ω, calculates the total. A label tied to superficial features of the problem could influence this learner to form a less-general subgoal.

Most students in the present study had at least two college-level math courses, thus it is reasonable that when cued with a label they could determine that a set of steps calculated a total. No Label students were less likely to form the subgoal of finding a total, presumably because it was more difficult for them to parse the solution procedure effectively. Thus, an interesting issue to examine in future work would be the effects of the learners' background and the labels used in examples on subgoal learning and transfer performance.

Other Types of Labels

While learners receiving the less-meaningful label were more likely to form representations free of erroneous superficial features, it is interesting to consider what the nature of the representations might be if the labels were meaningful, but unconnected to superficial details of the examples. For instance, instead of containing the label "total number of briefcases owned," suppose an example contained the label "total frequency of the event." This label is formally correct and not related to superficial details of the example. Perhaps this sort of label would produce the best transfer since it would supply a formally correct concept, presumably cue the grouping and self-explanation processes, and not provide a misleading tie-in to superficial features. On the other hand, it is possible that this sort of label, because it has some meaning to the learner (at least relative to Ω), but is presumably also a relatively unfamiliar concept, could distract the learner and compete with the grouping and self-explanation processes. This is an interesting empirical question that could be examined in future work and may be related to Sweller and Cooper's (1985) conjecture that schemas are more likely to be learned when other processes do not place an additional, distracting or concurrent load on working memory (see also Ward & Sweller, 1990).

Is There a "Right" or "Basic" Level of Subgoal Learning?

A solution procedure can be broken down into any number of subgoals and methods. It is unlikely that, with respect to improving transfer, a particular breakdown could be defended in any formal way either by an appeal to a particular cognitive architecture or to a task analysis unless one restricts the set of problems to be considered in the domain so that the pool of potentially useful subgoals can be identified by the researcher. Factors such as differences in learners' background and hypothesized working memory load as a function of a particular breakdown would also need to be considered, although both could potentially be addressed by a cognitive architecture that represents and tracks these factors.

Subgoals at varying levels of the solution structure hierarchy could affect transfer success to different degrees. For instance, in Experiment 2, transfer to novel problems was not predicted by whether or not participants mentioned the goal of finding λ in their descriptions; rather, it was associated with mentioning the goal of finding a total. This may be because the goal to find λ is too high-level; it does not sufficiently constrain the learner's search for appropriate steps to modify or create. It is intriguing to consider whether the inconsistency in the problem solving literature concerning difficulties in procedural transfer may potentially be explained through a subgoal "level" analysis.

Future Work

Besides exploring some of the issues raised above, an additional extension to the present work would be to examine the link between grouping and self-explanation more directly than was done in the present study. Approaches such as encouraging learners to talk out loud while studying examples would provide a more direct test of the link between grouping and self-explanations. This approach would also allow an examination of whether factors other than labels can affect grouping and whether grouping reliably leads to subgoal formation.

Another extension focuses on whether subgoals can be effectively communicated through declarative statements or whether they are best learned through examples. Chi et al. (1989) suggested that learners acquire knowledge through a self-explanation process that they might not be able to acquire as effectively otherwise. Examples might provide the best environment for learners to engage in the self-explanation process as well to integrate the resulting knowledge with the domain-relevant knowledge they already possess (Chi, de Leeuw, Chiu, & LaVancher, 1994). Thus, it would be interesting, both pedagogically as well as for the development of models of learning, to examine how effectively subgoals are learned through examples versus declarative text.

References

Anzai, Y., & Simon, H.A. (1979). The theory of learning by doing. Psychological Review, 86 (2), 124-140.

Ausubel, D.P. (1968). Educational psychology: A cognitive view. New York: Holt, Rinehart and Winston, Inc.

Brown, A.L., Kane, M.J., & Echols, C.H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. Cognitive Development, 1, 103-121.

Cabrera, A., & Billman, D. (1994). Language-driven concept learning: Deciphering "Jabberwocky." Manuscript submitted for publication.

Card, S.K., Moran, T.P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Erlbaum.

Catrambone, R. (in press). Aiding subgoal learning: Effects on transfer. Journal of Educational Psychology.

Catrambone, R. (1994). Improving examples to improve transfer to novel problems. Memory & Cognition, 22 (5), 606-615.

Catrambone, R., & Holyoak, K.J. (1990). Learning subgoals and methods for solving probability problems. Memory & Cognition, 18 (6), 593-603.

Chi, M.T.H., Bassok, M., Lewis, R., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Chi, M.T.H., de Leeuw, N., Chiu, M, & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.

Dixon, P. (1987). The processing of organizational and component step information in written directions. Journal of Memory and Language, 26 (1), 24-35.

Dufresne, R.J., Gerace, W.J., Hardiman, P.T., & Mestre, J.P. (1992). Constraining novices to perform expertlike problem analyses: Effects on schema acquisition. The Journal of the Learning Sciences, 2 (3), 307-331.

Eylon, B., & Reif, F. (1984). Effects of knowledge organization on task performance. Cognition and Instruction, 1 (1), 5-44.

Fong, G.T., Krantz, D.H., & Nisbett, R.E. (1986). The effects of statistical training on thinking about everyday problems. Cognitive Psychology, 18, 253-292.

Fried, L.S., & Holyoak, K.J. (1984). Induction of category distributions: A framework for classification learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10 (2), 234-257.

Heller, J.I., & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. Cognition and Instruction, 1 (2), 177-216.

Homa, D., & Cultice, J. (1984). Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10 (1), 83-94.

Kieras, D.E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. Cognitive Science, 8, 255-273.

LeFevre, J., & Dixon, P. (1986). Do written instructions need examples? Cognition and Instruction, 3, 1-30.

Luchins, A.S. (1942). Mechanization in problem solving. Psychological Monographs, 54 (Whole No. 248).

Medin, D.L., Wattenmaker, W.D., & Hampson, S.E. (1987). Family resemblance, conceptual cohesiveness, and category construction. Cognitive Psychology, 19, 242-279.

McDaniel, M.A., & Schlager, M.S. (1990). Discovery learning and transfer of problem solving skills. Cognition and Instruction, 7 (2), 129-159.

Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H.A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Novick, L.R., & Holyoak, K.J. (1991). Mathematical problem solving by analogy. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17 (3), 398-415.

Pirolli, P.L., & Anderson, J.R. (1985). The role of learning from examples in the acquisition of recursive programming skill. Canadian Journal of Psychology, 39 (2), 240-272.

Reed, S.K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11 (1), 106-125.

Ross, B. (1987). This is like that: The use of earlier problems and the separation of similarity effects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13 (4), 629-639.

Ross, B. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15 (3), 456-468.

Shaffer, J.P. (1986). Modified sequentially rejective multiple test procedures. Journal of the American Statistical Association, 81, 826-831.

Singley, M.K., & Anderson, J.R. (1989). The transfer of cognitive skill. Cambridge, MA: Harvard University Press.

Smith, E.E., & Goodman, L. (1984). Understanding written instructions: The role of an explanatory schema. Cognition and Instruction, 1 (4), 359-396.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12, 257-285.

Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. Cognition and Instruction, 2 (1), 59-89.

VanLehn, K. (1988). Toward a theory of impasse-driven learning. In H. Mandl & A. Lesgold (Eds.), Learning issues for intelligent tutoring systems (pp. 19-41). New York: Springer.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. Cognition and Instruction, 7 (1), 1-39.

Wattenmaker, W.D., Dewey, G.I., Murphy, T.D., & Medin, D.L., (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. Cognitive Psychology, 18, 158-194.

Zhu, X., & Simon, H.A. (1987). Learning mathematics from examples and by doing. Cognition and Instruction, 4 (3), 137-166.

Footnotes

[1]The Poisson distribution is often used to approximate binomial probabilities for events occurring with some small probability. The Poisson equation is $P(X=x) = \frac{[(e^{-\lambda})(\lambda^{x})]}{x!}$, where λ is the average (the expected value) of the random variable X.

[2]It is possible that the segmentation task could affect performance on the description task by leading participants to pay more attention to the steps than they might otherwise. However, a prior study (Catrambone, in press) found a relationship between a similar labeling manipulation and description performance without the intervening segmentation task.

[3]The most formal view would be "total frequency of the event."

Table 1

Training Example with Meaningful, Less-Meaningful, and No Label Solutions

A judge noticed that some of the 219 lawyers at City Hall owned more than one briefcase. She counted the number of briefcases each lawyer owned and found that 180 of the lawyers owned exactly 1 briefcase, 17 owned 2 briefcases, 13 owned 3 briefcases, and 9 owned 4 briefcases. Use the Poisson distribution to determine the probability of a randomly chosen lawyer at City Hall owning exactly two briefcases.

a.) No Label Solution:

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{289}{219}$$

$$= 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X{=}x) = \frac{\left[\left(e^{-\lambda}\right)\left(\lambda^{x}\right)\right]}{x!}$$

$$P(X{=}2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

b.) Meaningful Label Solution:

Total number of briefcases owned = [1(180) + 2(17) + 3(13) + 4(9)] = 289

$$E(X) = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X{=}x) = \frac{\left[\left(e^{-\lambda}\right)\left(\lambda^{x}\right)\right]}{x!}$$

$$P(X{=}2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

c.) Less-Meaningful Label Solution:

$$\Omega = [1(180) + 2(17) + 3(13) + 4(9)] = 289$$

$$E(X) = \frac{289}{219} = 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X{=}x) = \frac{\left[\left(e^{-\lambda}\right)\left(\lambda^{x}\right)\right]}{x!}$$

$$P(X{=}2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

Table 2

Sample Test Problems

a.) Total Frequency Calculated by Adding Simple Frequencies

Over the course of the summer, a group of 5 kids used to walk along the beach each day collecting seashells. We know that on Day 1 Joe found 4 shells, on Day 2 Sue found 2 shells, on Day 3 Mary found 5 shells, on Day 4 Roger found 3 shells, and on Day 5 Bill found 6 shells. Use the Poisson distribution to determine the probability of a randomly chosen kid finding 3 shells on a particular day.

Solution (not seen by participants):

$$E(X) = \frac{4+2+5+3+6}{5} = \frac{20}{5} = 4.0 = \lambda = \text{average number of shells per kid}$$

$$P(X=3) = \frac{[(2.718^{-4.0})(4.0^3)]}{3!} = \frac{(.018)(64)}{6} = .195$$

b.) Total Frequency Provided Directly

A number of celebrities were asked how many commercials they made over the last year. The 20 celebrities made a total of 71 commercials. Use the Poisson distribution to determine the probability that a randomly chosen celebrity made exactly 5 commercials.

Solution (not seen by participants):

$$E(X) = \frac{71}{20} = 3.55 = \lambda = \text{average number of commercials per celebrity}$$

$$P(X=5) = \frac{[(2.718^{-3.55})(3.55^5)]}{5!} = \frac{(.029)(563.8)}{120} = .135$$

Table 3

Solution Seen By Meaningful and Less-Meaningful Label Groups to One of the Training Examples (Experiment 2)

Meaningful Label Group:

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{\text{total number of briefcases owned}}{219} = \frac{289}{219}$$

$$= 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X{=}x) = \frac{\left[\left(e^{-\lambda}\right)\left(\lambda^{x}\right)\right]}{x!}$$

$$P(X{=}2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

Less-Meaningful Label Group:

$$E(X) = \frac{1(180) + 2(17) + 3(13) + 4(9)}{219} = \frac{\Omega}{219} = \frac{289}{219}$$

$$= 1.32 = \lambda = \text{average number of briefcases owned per lawyer}$$

$$P(X{=}x) = \frac{\left[\left(e^{-\lambda}\right)\left(\lambda^{x}\right)\right]}{x!}$$

$$P(X{=}2) = \frac{[(2.718^{-1.32})(1.32^{2})]}{2!} = \frac{(.27)(1.74)}{2} = .235$$

Table 4

Segmenting and Descriptions as a Function of Group (Experiment 2)

	Group		
	Meaningful Label (N = 30)	Less-Meaningful Label (N = 30)	No Label (N = 30)
Percentage Circling Total Frequency	63	67	37
Percentage Mentioning Finding			
λ	67	60	70
Total	57	40	20
Ω	0	27	0
Neither	43	33	80

Table 5

Sample Test Problems (Experiment 3)

a.) Weighted-Average with Humans Providing Total Frequency

The 17 aprons at a large restaurant were each worn one or more times by the various chefs. Three of the aprons were each used by exactly 1 chef, 7 of the aprons were used by 2 of the chefs, 4 of the aprons were used by 3 of the chefs, 2 of the aprons were used by 4 of the chefs, and 1 of the aprons was used by 5 of the chefs. Use the Poisson distribution to determine the probability of a randomly chosen apron being worn by exactly 2 chefs.

Solution (not seen by participants):

$$E(X) = \frac{1(3) + 2(7) + 3(4) + 4(2) + 5(1)}{17} = \frac{42}{17} = 2.47 = \lambda = \text{avg number of chefs per apron}$$

$$P(X{=}2) = \frac{[(2.718^{-2.47})(2.47^{2})]}{2!} = \frac{(.085)(6.1)}{2} = .26$$

b.) Total-Frequency-Given-Directly with Humans Providing Total Frequency

Over a period of time at a certain video store, 243 people rented 104 different videos. Use the Poisson distribution to determine the probability that a randomly chosen video was rented exactly 4 times.

Solution (not seen by participants):

$$E(X) = \frac{243}{104} = 2.34 = \lambda = \text{average number of renters per video}$$

$$P(X{=}4) = \frac{[(2.718^{-2.34})(2.34^{4})]}{4!} = \frac{(.096)(29.98)}{24} = .12$$

(Table 5 continued on the next page)

Table 5 (con't)

c.) Simple-Frequency with Humans Providing Total Frequency

An accounting firm employing many accountants worked on a large number of tax returns and used many types of tax forms. Four of the accountants were interviewed and it was found that one worked on 3 tax returns that day, another worked on 9, a third worked on 5, and the fourth worked on 6. In addition, of the many different types of tax forms used, it was found that one type of tax form was used by 12 accountants at the firm, another type was used by 8 accountants, a third type was used by 6 accountants, and a fourth type was used by 9 accountants. Use the Poisson distribution to determine the probability of a randomly chosen type of tax form being worked on by 7 different accountants.

Solution (not seen by participants):

$$E(X) = \frac{12 + 8 + 6 + 9}{4} = \frac{35}{4} = 8.75 = \lambda = \text{average number of accountants per form}$$

$$P(X=7) = \frac{[(2.718^{-8.75})(8.75^{7})]}{7!} = \frac{(.00016)(3926960)}{5040} = .125$$

Table 6

Participants' Scores on Novel Test Problems as a Function of Group and Correspondence (Experiment 3)

	Group			
	Meaningful Label (N = 30)	Less-Meaningful Label (N = 30)	No Label (N = 30)	AVG
Same Correspondence	3.07	3.00	1.53	2.53
Reversed Correspondence	1.67	2.67	1.13	1.82
AVG	2.37	2.83	1.33	2.18

Note. Maximum possible score for any cell = 4.

Distribution List

1) Defense Technical Information Center
   Building #5, Cameron Station
   Alexandria, VA 22304-6145

2) Dr. Susan E. Chipman
   Office of Naval Research
   Code 342 CS
   Ballston Tower One
   800 North Quincy Street
   Arlington, VA 22217-5660

3) NRL

4) TBD